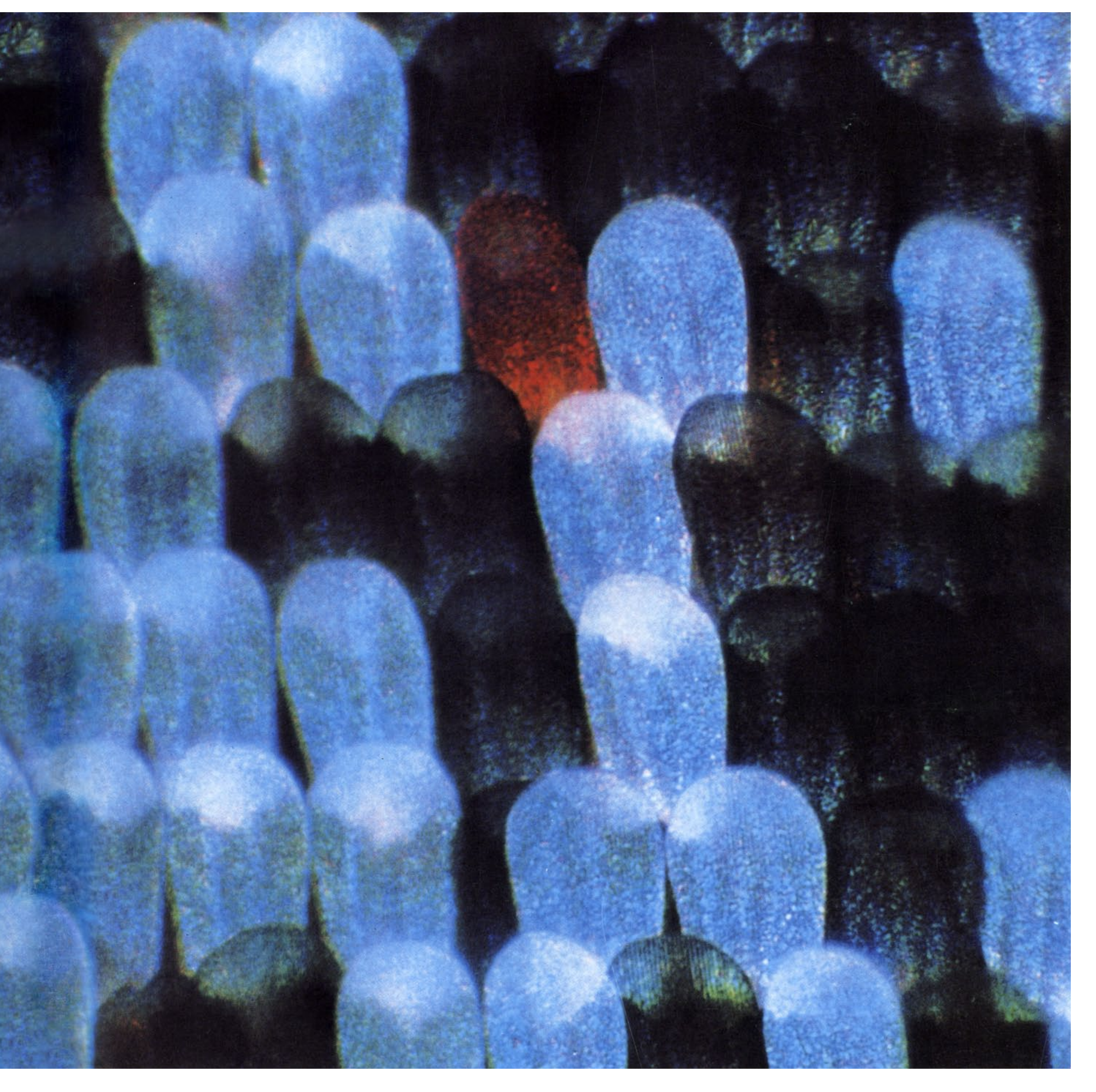
INVESTIGACION Y CIENCIA

Edición en español de SCIENTIFIC AMERICAN



PATRONES DE COLORACION EN MARIPOSAS

Enero 1982
300 PTAS.

Copyright © 1982 Prensa Científica S.A.

Los espacios en gris
corresponden a publicidad
en la edición impresa

Número 64
Enero 1982

INVESTIGACION Y
CIENCIA
Edición en español de SCIENTIFIC AMERICAN

SCIENTIFIC AMERICAN

COMITE DE REDACCION: Gerard Piel (Presidente), Dennis Flanagan, Brian P. Hayes, Philip Morrison, Francis Bello, Peter G. Brown, Michael Feirtag, Paul W. Hoffman, Jonathan B. Piel, John Purcell, James T. Rogers, Armand Schwab, Jr., Joseph Wisnovsky

DIRECCION EDITORIAL: Dennis Flanagan
DIRECCION ARTISTICA: Samuel L. Howard
PRODUCCION: Richard Sasso
DIRECTOR GENERAL: George S. Conn

INVESTIGACION Y CIENCIA

DIRECTOR: Francisco Gracia Guillén
REDACCION: José María Valderas Gallardo (Redactor Jefe), Carlos Oppenheimer, José María Farré Josa
PRODUCCION: César Redondo Zayas
VENTAS Y PUBLICIDAD: Elena Sánchez-Fabrés
PROMOCION EXTERIOR: Pedro Clotas Cierco
EDITA: Prensa Científica, S. A.
Calabria, 235-239
Barcelona-29 (ESPAÑA)

Colaboradores de este número:

Asesoramiento y traducción:

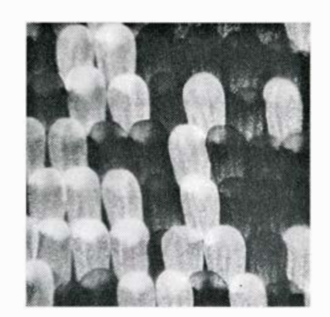
José A. Martín Pereda y M. Fernández Centeno: *Intensificador de imágenes por microcanales*; Santiago Torres y Rosa Ruiz Vázquez: *Reparación inducible del ADN*; Manuel Puigcerver: *Anillos en el sistema solar*; Amando García Rodríguez: *Las fases fluidas de la materia*; Jaume Baguñá Monjo: *Patrones de coloración de las mariposas diurnas y nocturnas*; Cristóbal Mezquita Pla: *El revestimiento del intestino delgado*; Laureano Carbonell Relat: *Establecimiento ballenero vasco del siglo XVI en el Labrador*; Luis Bou: *Temas metamágicos*; J. Vilardell: *Taller y laboratorio.*

Ciencia y sociedad:

Julio Samsó, F. X. Avilés Puigvert y Antonio Blanco Freijeiro

Libros:

Javier de Lorenzo, Ramón Queraltó y Francisco García Novo



LA PORTADA

La fotografía de la portada recoge unas escamas del ala de mariposa, aumentadas unos 280 diámetros. Son protrusiones aplanadas de las células alares. Cada escama corresponde a una célula distinta. El conjunto de ellas, que se superponen como tejas, a modo de mosaico, forma el patrón de coloración del ala (véase "Patrones de coloración de las mariposas diurnas y nocturnas" por H. Frederik Nijhout, página 70). Se aprecia aquí parte del patrón de la superficie inferior del ala de la mariposa *Limenitis astyanax*. Las escamas azules deben su color a la interferencia de la luz reflejada. Se trata de un color estructural, esto es, generado por las características estructurales de las escamas. El color de la rojiza es pigmentario, y resulta de la melanina que hay en la escama. Las escamas pigmentadas sólo contienen un pigmento, pero varía su cantidad. Un patrón típico no tiene más de cinco colores, y lo común son tres o cuatro.

Suscripciones:
Prensa Científica, S. A.
Calabria, 235-239
Barcelona-29 (España)
Teléfono 322 05 51 ext. 37

Condiciones de suscripción:

España:
Un año (12 números): 3.300 pesetas

Extranjero:
Un año (12 números): 52 U.S. $

Ejemplar atrasado ordinario:
340 pesetas

Ejemplar atrasado extraordinario:
440 pesetas

Distribución para España
Distribuciones de Enlace, S. A.
Ausias March, 49, Barcelona-10

Distribución para los restantes países:
Editorial Labor, S. A.
Calabria, 235-239 - Barcelona-29

Publicidad:
Madrid:
Gustavo Martínez Ovin
Avda. de Moratalaz, 137, Madrid-30
Tel. 430 84 81

Cataluña:
Lourdes Padrós
Manuel Girona, 61, 2.°, Barna-34
Teléfono 204 45 83

Controlado por O.J.D.

OJD DIFUSION PUBLICIDAD

PROCEDENCIA DE LAS ILUSTRACIONES

Fotografía de la portada de Fritz Goro

Página	Fuente	Página	Fuente
8	Ralph Morse	77	Tom Prentiss
10-12	George V. Kelvin	78	H. Frederik Nijhout
14	Ralph Morse	79	Tom Prentiss
15	Adolf R. Asam, ITT Electro-Optical Products Division, Roanoke, Virginia (*arriba*); Michael Lampton (*abajo*)	80	H. Frederik Nijhout y Tom Prentiss
		83	Rob McLachlin (*arriba*), Peter Waddell (*abajo*)
		84-85	Universidad Memorial de Terranova
16-17	George V. Kelvin	86	Andrew Tomko
18	Biblioteca Vaticana	87	Lester Ross, Parks Canada (*arriba*); J. Martin, Universidad Memorial de Terranova (*abajo*)
20	Museo Capitolino, Roma		
21	Museo del Pireo		
22	Miguel Alonso		
23	Biblioteca Central, Barcelona		
24	Biblioteca Nacional, Bolonia		
25	José Alsina y Miguel Alonso	88	J. Martin, Universidad Memorial de Terranova
26	Monasterio de El Escorial		
27	Biblioteca Laurenziana, Florencia		
29	Paul Howard-Flanders	89	Canadian Conservation Institute
30-37	Patricia J. Wynne		
44	Jet Propulsion Laboratory	90	Universidad Memorial de Terranova
46-47	Walken Graphics		
48	Jet Propulsion Laboratory	93	Florence Moog
50	Walken Graphics	94-96	Carol Donner
51-52	Jet Propulsion Laboratory	97	Florence Moog
54	Walken Graphics	98	Susumu Ito
55	Jet Propulsion Laboratory	99	Andrew Christie
56-57	Walken Graphics	100	Dominique Maestracci
71	Fritz Goro	101-102	Andrew Christie
72-73	Tom Prentiss	105-112	Mitchell J. Feigenbaum
74	H. Frederik Nijhout	115	Kenneth A. Connors
76	H. Frederik Nijhout (*izquierda*), Tom Prentiss (*derecha*)	116	John M. Franke
		117-119	Michael Goodman
		120	Adam Lloyd Cohen

ISSN 0210-136X
Dep. legal: B. 38.999-76
Fotocomposición Tecfa
Pedro IV, 160 - Barcelona-5
Fotocromos reproducidos por GINSA, S.A.
Imprime GRAFESA
Gráfica Elzeviriana, S. A.
Nápoles, 249 - Tel. 207 40 11
Barcelona-13
Printed in Spain - Impreso en España

Copyright © 1981 Scientific American Inc., 415 Madison Av., New York, N. Y. 10017.
Copyright © 1982 Prensa Científica, S. A., Calabria, 235-239 - Barcelona-29 (España)

Reservados todos los derechos. Prohibida la reproducción en todo o en parte por ningún medio mecánico, fotográfico o electrónico, así como cualquier clase de copia, reproducción, registro o transmisión para uso público o privado, sin la previa autorización escrita del editor de la revista.

El nombre y la marca comercial SCIENTIFIC AMERICAN, así como el logotipo distintivo correspondiente, son propiedad exclusiva de Scientific American, Inc., con cuya licencia se utilizan aquí.

Los autores

MICHAEL LAMPTON ("Intensificador de imágenes por microcanales") trabaja en el Laboratorio de Ciencias Espaciales de la Universidad de California en Berkeley. Es especialista en entrenamiento de vuelo en la misión del *Spacelab 1*. Se licenció en ciencias físicas en el Instituto de Tecnología de California en 1962, doctorándose en Berkeley cinco años más tarde. Ha participado en varias misiones espaciales, incluido el acoplamiento Apolo-Soyuz en 1975. Ocupa su tiempo libre en el diseño de altavoces.

JOSE ALSINA CLOTA ("Hipócrates") es catedrático de filología griega de la Universidad de Barcelona desde 1958. Allí se formó y ha ocupado diversos cargos de responsabilidad académica. Autor de numerosos artículos sobre mitología, religión, filosofía, literatura y ciencia antigua, ha sido también editor y traductor de autores griegos. Pertenece al comité de redacción de las revistas *Estudios clásicos* y *Studia papyrologica*. Dirige el *Boletín del Instituto de Estudios helénicos*, fundado por él.

P. HOWARD-FLANDERS ("Reparación inducible del ADN") es profesor del departamento de bioquímica y biofísica molecular y del departamento de radiología terapéutica de la Facultad de Medicina de la Universidad de Yale. Estudió física y matemática en el Colegio Imperial de Ciencia y Tecnología, en su Inglaterra natal. Entre los primeros trabajos de Howard-Flanders se cuenta el diseño del primer acelerador lineal de microondas empleado en radioterapia. En 1958 enseñó biofísica en la Universidad de California en Berkeley, trasladándose a Yale el año siguiente. Su investigación se centra en el estudio de los mecanismos celulares de reparación del ADN dañado; es codescubridor de la reparación por excisión y de la reparación post-replicativa.

JAMES B. POLLACK y JEFFREY N. CUZZI ("Anillos en el sistema solar") se hallan adscritos al Centro de Investigación Ames de la Administración Nacional de Aeronáutica y del Espacio (NASA). Pollack ocupa puesto directivo en el Programa de Efectos Climáticos de los Aerosoles, de la NASA. Se formó en la Universidad de Princeton y en la de California en Berkeley, si bien el doctorado lo obtuvo en Harvard en 1965. Cinco años más tarde ingresó en la NASA, donde ha venido participando en varias misiones: Pioneer Venus, *Mariner 9*, Viking a Marte y Voyager a Saturno. Dedicación y competencia que le han hecho acreedor a dos medallas de la NASA y un premio del Instituto Americano de Aeronáutica y Astronáutica. Cuzzi inició sus estudios en la Universidad de Cornell y los prosiguió en el Instituto de Tecnología de California, donde se doctoró en 1972. Antes de ingresar en Ames, lo que ocurrió en 1974, desempeñó cargos de investigación en la Universidad de Massachusetts y en Grand Ronde.

J. A. BARKER y DOUGLAS HENDERSON ("Las fases fluidas de la materia") vienen trabajando, desde 1969, en el laboratorio de investigación que la International Business Machines tiene en S. José, California. Barker se graduó en física, matemática y química, con el título acreditativo de cada carrera, en la Universidad de Melbourne. Desde 1950 hasta 1967 estuvo en los Chemical Research Laboratories de la Commonwealth Scientific and Industrial Research Organization (CSIRO), de Melbourne. Un año antes de su incorporación a la IBM enseñó física y matemáticas aplicadas en la Universidad de Waterloo, Ontario (Canadá). Es miembro de la Australian Academy of Science y de la Royal Society de Londres. "Mis campos de investigación, dice, han sido la mecánica estadística, el estudio de las fuerzas intermoleculares y la teoría de los sólidos, líquidos, gases y superficies." Dentro de estos campos, Barker ha insistido en "el desarrollo de técnicas de cálculo y de métodos de simulación con ordenadores". Henderson es canadiense. Se licenció en matemáticas por la Universidad de Columbia Británica, en 1956, doctorándose en física por la de Utah en 1961. Al igual que Barker, enseñó física y matemáticas aplicadas en la Universidad de Waterloo (entre 1964 y 1969) y trabajó en los Chemical Research Laboratories de la CSIRO, donde "se formularon muchas de las ideas que contiene este artículo". Miembro del Instituto de Física del Reino Unido, de la Sociedad Norteamericana de Física y del Instituto Norteamericano de Química, Henderson se interesa por "la mecánica estadística y, en particular, por la teoría de los líquidos y de las interfases líquidas, así como por la teoría de los sólidos desordenados".

H. FREDERIK NIJHOUT ("Patrones de coloración de las mariposas diurnas y nocturnas") es profesor de zoología en la Universidad Duke. Nacido en Holanda, se licenció en biología por la Universidad de Notre-Dame, en 1970, doctorándose por Harvard en 1974. Antes de incorporarse a la Universidad Duke, trabajó como postdoctorando en la de Washington en Seattle y como investigador en el Instituto Nacional de Alergia y Enfermedades Infecciosas. "Mi trabajo en patrones de coloración, nos comenta, constituye sólo un aspecto de mi interés general por la biología del desarrollo de los insectos. Me atrae también la fisiología de la muda y la metamorfosis, sobre todo aquellos procesos que regulan la secreción de las hormonas del desarrollo de los insectos."

JAMES A. TUCK y ROBERT GRENIER ("Establecimiento ballenero vasco del siglo XVI en el Labrador") son, respectivamente, profesor de arqueología de la Universidad Memorial de Terranova y director de la Unidad de Excavaciones Marinas de Parques de Canadá. Tuck se doctoró en la Universidad de Siracusa, en 1968, e inmediatamente después entró a trabajar en la Memorial. Ha participado en varias excavaciones, principalmente en Terranova y Labrador. Grenier estudió en la Universidad de Laval, de Quebec, y su bautizo de campo fue la excavación de la fortaleza de Louisbourg, en la isla de Cabo Bretón.

FLORENCE MOOG ("El revestimiento del intestino delgado") enseña biología en la Cátedra Charles Rebstock de la Universidad de Washington. Se licenció por la Universidad de Nueva York en 1936. En 1944 se doctoró por la de Columbia. En ésta, escribe, "me sentí atraída por la embriología experimental, ciencia que parecía estar a punto de resolver algunos de los problemas más profundos de la biología (que siguen pendientes todavía en 1982). Todas mis investigaciones las he realizado en el campo del desarrollo embriológico, particularmente el de los vertebrados. En los últimos años de la década de los 40 dirigí mi atención al intestino del embrión y, sin pretenderlo, fui pionera en lo que luego constituyó un área muy activa de investigación, el intestino fetal y del lactante. Respecto a mis aficiones no científicas, soy una entusiasta de la ópera, y una ávida lectora de historia, biografía y novela, interesándome especialmente la obra de Anthony Trollope".

Hace...

José M.ª López Piñero

... *trescientos años*

José Lucas Casalete, catedrático de Prima de Medicina en la Universidad de Zaragoza y cabeza del grupo de *novatores* de dicha ciudad, fue objeto de una de las campañas condenatorias más duras que tuvieron que sufrir los responsables de la introducción de las corrientes científicas modernas en la España de finales del siglo XVII.

El ataque partió de Nicolás Moneva, Visitador Médico del Reino de Aragón, quien, en sus viajes de inspección, "reconoció algunos abusos, nacidos de las proposiciones" de Casalete y, "por atención a la salud pública", lo denunció a los claustros médicos de las principales universidades. La respuesta a dicha denuncia constituye un expresivo ejemplo de cómo el profesorado universitario puede convertirse en un grupo cerrado y paralizador, ya que los catedráticos de medicina de Salamanca, Alcalá, Valladolid, Valencia, Barcelona, Lérida y Huesca condenaron la enseñanza del colega innovador con las expresiones más duras. Las teorías de Casalete fueron calificadas de "falsas, erróneas, temerarias, perniciosas a la salud pública, indignas de tan grave autor, irracionales, absurdas", etc., por ser "opuestas a la doctrina de Galeno". Todos esos claustros acordaron unánimemente "que no se debe permitir que se enseñen, ni en voz ni en escritos".

El antecesor de Casalete en la cátedra de Prima había sido Matías de Llera, cuya obra más importante, titulada de forma muy barroca *Manus medica dextera quinque digitos continens* (1666), incluye cinco tratados pertenecientes todos ellos a la literatura típica del galenismo más ortodoxo. En forma de comentario de los textos de Galeno, se ocupan de las fiebres, de la utilización terapéutica de la sangría y los purgantes, de las crisis y días críticos, y del modo de conocer las enfermedades. La mentalidad de Llera era absolutamente tradicional y no dudó, por ejemplo, en defender la utilidad de la sangría como remedio en casi todas las enfermedades "cum materia". Aunque tuvo gran capacidad como observador clínico, la posición conservadora de Llera contrasta violentamente con la actividad de Casalete, único *novator* que fue titular de una cátedra universitaria.

¿Cómo fue posible este cambio? No cabe duda que la explicación hay que referirla a las circunstancias generales del ambiente científico aragonés de estos años. Conviene no olvidar, en primer término, la presencia en Zaragoza de una gran figura como Juan Bautista Juanini, que contribuyó decisivamente a la difusión de las ideas modernas. Sabemos, por ejemplo, que en 1676 dio a conocer entre los médicos al servicio de Juan José de Austria –uno de los cuales era el propio Matías de Llera y otro el protomédico Tomás de Longás, al que después nos referiremos– las obras del iatroquímico francés François Bayle, entonces recién aparecidas. Juanini también realizó en Zaragoza autopsias y experiencias demostrativas de la circulación de la sangre, ante el mismo Juan José de Austria y en el anfiteatro anatómico del Hospital de Nuestra Señora de Gracia. Todo ello debió influir en una Universidad que, a semejanza de las de Valencia, Barcelona y Salamanca, supo mantener a un notable nivel la docencia práctica de la anatomía.

La doctrina de la circulación sanguínea fue, en efecto, introducida en la enseñanza por Francisco San Juan y Campos, que ocupó la cátedra de Anatomía en 1686. Un testimonio contemporáneo del médico italiano Federico Bottoni resulta a este respecto indispensable: "En la célebre Universidad de Zaragoza, corte del Reino de Aragón, propugnó y estableció esta doctrina que quedó establecida como principio elemental, subiendo a mi intento, de mucha autoridad, esta aceptación, pues nadie de los europeos ignora que en este celebrado museo de las ciencias florece la medicina en el más elevado crédito, debiéndose éste al continuo ejercicio anatómico que dos veces a la semana se ejecuta en el teatro o salón que para este efecto hay en aquel célebre Hospital General, concurriendo todos los profesores de esta ciencia a tan importante demostración."

Como hemos adelantado, la ruptura abierta con los esquemas tradicionales en esta Universidad la encabezó, sin embargo, José Lucas Casalete. A partir de 1653 ocupó diversas cátedras médicas y en 1677 sucedió a Llera en la de Prima. Entonces había ya demostrado su actitud inconformista, prologando tres años antes el libro del curandero Juan de Vidós y Miró, motivo de un ruidoso pleito con el Colegio de San Cosme y San Damián. Mantuvo asimismo relación amistosa con otros protagonistas de la renovación científica española, entre ellos el citado Juan Bautista Juanini y Juan de Cabriada, para algunas de cuyas obras escribió elogiosos prefacios. Al hacerse cargo de la cátedra, Casalete se enfrentó con el galenismo hasta entonces vigente, no sólo en el terreno doctrinal sino también en el práctico. Defendió en concreto un nuevo método curativo de las fiebres que suponía un fuerte ataque al abuso de la sangría por parte de los galenistas. Las críticas contra dicho abuso tenían una larga tradición, que había sido actualizada, entre otros autores, por Gonzalo Bustos de Olmedilla en su libro *El monstruo horrible de la Grecia* (1669). A pesar de que los contradictores de Casalete intentaron unir su nombre al de Olmedilla, la postura del catedrático de Zaragoza se diferenciaba radicalmente de la de Bustos por su apoyo directo en una base doctrinal moderna incompatible con los esquemas galénicos. Esta base consistía fundamentalmente en una combinación de la iatroquímica con el solidismo procedente del médico italiano Prospero Alpino. Frente a la teoría galénica de la "fluxión humoral", Casalete pensaba que la causa de las fiebres era una alteración patológica de la parte sólida de los órganos, que conducía al viciamiento de los jugos orgánicos. Las fiebres serían el resultado directo de la acción de estos jugos alterados, de naturaleza ácida o nitrosa. Las inflamaciones y las colecciones sanguíneas no se producirían en un lugar determinado por la "fluxión" de los galenistas, sino porque dichos jugos coagulan la sangre y producen una detención en la circulación sanguínea. El concepto galénico de "plenitudo" o plétora quedaba de este modo interpretado desde un ángulo hemodinámico moderno, mientras que el mecanismo de producción de la enfermedad se explicaba mediante la iatroquímica. Por otra parte, el interés por el "foco" de la enfermedad en la parte sólida de los órganos abría la puerta a las indagaciones anatomopatológicas.

Cuatro años más tarde, en 1686, las ideas de Casalete fueron criticadas, en un tono muy distinto al utilizado por los claustros médicos, por Nicolás Francisco San Juan y Domingo, un antiguo discípulo suyo. La crítica apareció en un capítulo del libro *De morbis endemiis Caesar-Augustae*, obra que merece atención en sí misma como expresión de la postura ideológica del galenismo moderado de estos años. Se trata de la primera topografía médica de enfoque moderno aparecida en nuestro país. Las llamadas "topografías médicas" constituyeron un género que alcanzó gran importancia durante la Ilustración y buena parte del siglo XIX. Inspirado en el tratado hipocrático *Del aire, las aguas y los lugares*, dicho género tenía

como fundamento el estudio del ambiente de una zona determinada con el fin de explicar la génesis de las enfermedades dominantes en la misma. Con esta intención, San Juan y Domingo describió en su libro la situación de Zaragoza, la geografía física de su zona, en especial el clima, las costumbres de sus habitantes, etcétera. El estudio de las enfermedades lo basó en las doctrinas galénicas, aunque aceptando, en una línea típicamente "moderada" o ecléctica, numerosos elementos de las corrientes modernas. La información que San Juan y Domingo disponía acerca de estas últimas era excelente: se refiere al *Sepulchretum* de Bonet al ocuparse de las lesiones anatómicas como base de las fiebres, acepta el uso de los medicamentos químicos y también varias doctrinas fisiológicas modernas, como la circulación de la sangre y la "perspiratio insensibilis" de Sanctorio. Por el contrario, se opone a la explicación iatroquímica de la patogenia de la fiebre y refuta en detalle la teoría de Silvio y también la de Sydenham. En este contexto critica el "método curativo" de Casalete al que, por otra parte, trata con gran estima y respeto. Desde el punto de vista doctrinal lo adscribe fundamentalmente al solidismo metódico e intenta, desde su misma posición, refutar sus ataques a la práctica de la sangría, con el argumento de que la mayor parte de las enfermedades de Zaragoza estaban causadas por una "plétora sobreestimulante" que requería un tratamiento antiflogístico.

Al año siguiente de la publicación de la obra de San Juan y Domingo –es decir, en la significativa fecha de 1687, habitualmente aceptada como hito en el movimiento *novator* español– aparecieron en Zaragoza dos textos fundamentales de la renovación científica en el área aragonesa. El primero fue la obra del propio Casalete titulada *Duae controversiae*, en la que resumía los principales puntos de su "nueva práctica". El segundo, el libro *Statera medicinae selectae*, del médico navarro Francisco Elcarte, el más importante de los discípulos y seguidores de Casalete. La primera y la tercera partes del libro de Elcarte están redactadas en latín y se dedican a la exposición y defensa de las doctrinas y del "método curativo" propuestos por el catedrático aragonés. La segunda parte, escrita en castellano, es una "censura" de la obra de San Juan y Domingo antes citada. Elcarte no sólo refuta los puntos de vista de este autor en lo relativo a la patogenia de las fiebres, sino que critica asimismo sus ideas acerca de la topografía médica. Considera muy simplistas sus explicaciones de base casi exclusivamente climática, y las contrapone a las expuestas por su maestro en sus lecciones de higiene. Casalete defendía, en efecto, un estudio más amplio de las condiciones del terreno, de sus aires y de sus aguas, es decir, un tipo de acercamiento directamente inspirado en el tratado hipocrático antes anotado y muy similar al que alcanzaría notable desarrollo durante la Ilustración. Sus ideas higiénicas las expuso más tarde en un informe titulado *Satisfacción precisa a una duda voluntaria* (1689), que redactó en cola-

22 LIBER PRIMVS

El Doctor Nicolàs Moneva, reconociò algunos abuſos, nacidos de eſtas propoſiciones, en las Viſitas, que hizo por el Reyno los años de 1682. y el de 1683. y por atencion à la ſalud publica, juzgò conveniente, comunicarlas con algunas Vniverſidades, para vèr ſu probabilidad: Y reſpondiò el Colegio de Medicos de Salamanca à 23. de Março de 1684. que dichas propoſiciones no ſe pueden leer en publico, ni en ſecreto, ni practicarſe con buena conciencia. Eſtàn firmados en dicha reſpueſta los Señores.

El Doctor Don Franciſco Angel de Eſpinoſa, jubilado en la de Prima, y Decano.
El Doctor Matheo la Parra, Cathedratico actual de Prima.
El Doctor Don Pedro Garcia, Cathedratico de Viſperas.
El Doctor Don Antonio Sanchez Nieto, Cathedratico de Partido.
El Doctor Don Ioſeph Colmenero, Cathedratico de Simples.
El Doctor Gabriel Iolij, Cathedratico de Cirugia.
El Doctor Don Alonſo Ruiz, Cathedratico de Methodo.
El Doctor Don Nicolàs Herrera, Cathedratico de Anothomia.
El Doctor Don Gabriel Iolij, Opoſitor à Cathedras de Artes, y Medicina.

De todo lo qual dà teſtimonio Sebaſtian Garcia de Paredes, Secretario de la Vniverſidad.

Los

DE NOVA PRAXI. 23

Los Doctores de Medicina Complutenſes, reſpondieron à 4. de Março de 1684. que dichas propoſiciones ſon agenas de razon, temerarias, y que ſe deven prohibir en la practica, por perniciofas. Los firmados ſon, los Señores.

El Doctor Don Franciſco de Rivas del Caſtillo.
El Doctor Don Iuan de Torralva, Cathedratico en la principal de Viſperas.
El Doctor Sebaſtian de Azotho, Cathedratico de Viſperas.
El Doctor Don Diego del Varco, Cathedratico de Anothomia.

Y haze fè de dichas firmas Lazaro Sanchez Notario.

Los Doctores de Medicina de Valladolid, reſpondieron à 10. de Iunio de 1684. que las dichas propoſiciones, como ſuenan, ſon improbables, y abſurdas, y contra los principios de Medicina, Theorica, y Practica. Los firmados ſon, los Señores.

Don Rodrigo de Meſsìa, Rector de la Vniverſidad.
El Doctor Don Geronimo Pardo, primario en la de Avicena.
El Doctor Don Antonio Ximenez, Medico de Camara de ſu Mageſtad.
El Doctor Don Antonio Requena, Medico de ſu Mageſtad, y jubilado en la de Prima.
El Doctor Don Iuan Gallardo, Primario de Philoſophia.
El Doctor Don Iuan de Bonaſilla, Medico de la Santa Inquiſicion.

El

Páginas de la reprobación de las teorías innovadoras de José Lucas Casalete por parte de los catedráticos de medicina de las principales universidades españolas de la época, que fue reproducida en el **Enchiridion** *(1689), de Tomás Longás*

boración con otros dos médicos, en respuesta a una consulta acerca de las consecuencias que la instalación de fábricas de tabacos dentro de las ciudades podía tener para la salud pública.

En la polémica en torno a Casalete intervino asimismo Tomás Longás, otro importante galenista "moderado". Fallecido en 1690, no hay que confundirlo con su hijo Tomás Longás y Pascual, titular de la cátedra de anatomía de la Universidad de Valencia desde 1698 a 1722, y uno de los defensores de la renovación científica en dicho centro en las primeras décadas del siglo XVIII. Longás padre era protomédico del Reino de Aragón y uno de los médicos de cámara de Juan José de Austria. Su relación con Juanini contribuye a explicar su apertura parcial a las novedades. Por Federico Bottoni sabemos, por ejemplo, que fue uno de los primeros médicos españoles que administró inyecciones intravenosas, así como que era un decidido defensor de la circulación de la sangre. En 1689 publicó un volumen titulado *Enchiridion novae et antiquae medicinae dogmaticae*, que refleja muy claramente su condición de científico tradicionalista "moderado". De las cuatro partes de que consta la obra, la de mayor interés es la primera, consagrada a la crítica del "método curativo" de Casalete. Longás reproduce la historia clínica de unas fiebres padecidas por un importante personaje, motivo del enfrentamiento de Casalete con San Juan y Domingo, el propio Longás y otros galenistas. Tanto en ella como en las demás partes del libro, defiende los esquemas galénicos, aunque aceptando como rectificaciones de detalle numerosas novedades. Destaca en especial su interés por la química, de la que llega a afirmar: "Sin los principios de los químicos hay muchas cosas casi inexplicables acerca de la naturaleza de las enfermedades, su curación y la investigación de los medicamentos." Sin embargo, no considera, como Casalete, que la química deba constituir el fundamento de una nueva patología iatroquímica o atomista, y la destina a enriquecer y consolidar las doctrinas galénicas, a cuya "praestantia" consagra todo un capítulo de su libro.

La posición renovadora de Casalete motivó también la publicación de folletos polémicos del tipo habitual en esta época y se reflejó en obras de importancia muy secundaria, como la *Llave del tesoro de la piedra filosofal de la salud humana* (1688), publicada por Miguel Palacio y Pérez, médico de Viana que había sido discípulo del catedrático zaragozano.

Intensificador de imágenes por microcanales

Es una oblea de vidrio, perforada por millones de tubos multiplicadores de electrones, que se asemeja a un ojo compuesto. Transforma una diminuta secuencia espacial de radiación electromagnética en una imagen brillante

Michael Lampton

Un amplificador es un dispositivo que incrementa la potencia de una señal sin apenas alterar sus demás características. La amplificación suele emparejarse sólo con señales variables en el tiempo; por ejemplo, la señal eléctrica que sale de un micrófono. Una escena visual constituye una señal que varía tanto en el espacio como en el tiempo; si se desea amplificarla es necesario, por lo general, transformarla en una muestra de características puramente temporales. Esto suele realizarse mediante un barrido sistemático de la escena, como hacen las cámaras de televisión. Sin embargo, a lo largo de los últimos 25 años, una serie de avances en la tecnología del trabajado del vidrio han hecho posible la manufactura de millones de amplificadores microscópicos, distribuidos de una forma geométrica regular. Cada amplificador en esta distribución es un tubo o canal, de alrededor de 15 micras de diámetro, capaz de aumentar el brillo de una pequeña zona, perfectamente delimitada, de una escena. En consecuencia, el conjunto de amplificadores, que actúan de forma simultánea y en paralelo, trabaja como un intensificador de imágenes, transformando imágenes tenues en brillantes y sin destruir la información espacial de la señal de entrada. Un conjunto de amplificadores como el aquí descrito se denomina placa de microcanales.

El incentivo económico para el desarrollo de esta placa de microcanales ha venido, en gran parte, del estamento militar. Un intensificador de imágenes, compacto, que la incorpore puede aumentar el brillo de una escena en un factor de hasta 10.000. Con ello se viabiliza la vigilancia nocturna bajo unas condiciones mucho más severas, por su oscuridad, que las soportables hasta la fecha. Gafas, prismáticos e instrumentos análogos que incorporen estos intensificadores de microcanales han llegado a convertirse en imprescindibles en operaciones militares nocturnas, en las que, por ejemplo, los vehículos han de conducirse sin ningún tipo de iluminación. El Laboratorio de Visión Nocturna del Ejército de los Estados Unidos, de Fort Belvoir, Virginia, ha sido, en buena parte, el motor del gran desarrollo de estos dispositivos. Esas aplicaciones absorben hoy la producción de unos 10.000 intensificadores de microcanales por año.

Pero el intensificador de microcanales tiene aplicaciones también de carácter civil. En astronomía puede realzar el brillo de una imagen hasta el punto de permitir detectar fotones aislados. El resultado es poder registrar en varios minutos imágenes que, por otros métodos, requerirían exposiciones prohibitivamente largas. Además, han hecho posible el que personas con retinitis pigmentosa, una variedad de ceguera nocturna, logren desenvolverse sin dificultad. El único problema es que el costo actual (alrededor de 2700 dólares) constituye una verdadera barrera para una más amplia difusión.

La placa de microcanales no se limita a la intensificación de imágenes visibles. Cuando se complementa con los adecuados sistemas eléctricos y ópticos, puede transformar una señal bidimensional en la región del infrarrojo cercano o del ultravioleta en una imagen visible. Asimismo, puede detectar directamente rayos X, iones y electrones, abriendo con ello un abanico de nuevas aplicaciones en áreas tales como el diseño de osciloscopios, microscopía electrónica y estudio de la composición química de materiales mediante métodos fotométricos. Así, se emplea una placa de microcanales en el telescopio de rayos X situado a bordo del Observatorio Einstein (un satélite artificial) y en los espectrómetros de ultravioleta transportados en las misiones del Voyager a Júpiter y Saturno. Otro satélite, el Observatorio de rayos X europeo (*EXOSAT*), que va a lanzar este año la Agencia Espacial Europea, llevará consigo placas de microcanales como parte de sus detectores de rayos X.

¿Qué principios gobiernan la forma de trabajo de una placa de microcanales? Quizá convenga abordar la respuesta con una descripción de un intensificador de imágenes que no incorpore una placa de microcanales. Un dispositivo de este tipo se basa en la interacción de una radiación electromagnética con la materia, fenómeno explicado por vez primera por Einstein en 1905. Esta interacción constituye el conocido efecto fotoeléctrico; fue uno de los primeros procesos físicos por los que se llegó a la idea de que la energía se encuentra en cuantos o cantidades discretas.

Cuando un fotocátodo, esto es, un electrodo cargado negativamente, se somete a una radiación electromagnética cuya longitud de onda es inferior a un determinado valor, emite una corriente de electrones que puede recogerse en un ánodo próximo, esto es, en un eletrodo cargado positivamente. Para muchos metales, esta longitud de onda crítica se encuentra en la región del ultravioleta del espectro electromagnético, aunque para algunas sustancias está en la región del visible o del infrarojo cercano. La velocidad a que se transfiere la carga es proporcional a la intensidad de la radiación. Pero el que exista o no transferencia de carga depende solamente de la longitud de onda.

Einstein explicó este fenómeno suponiendo que un electrón ligado a la materia requiere una determinada energía

mínima para ser liberado. Además, si esta energía proviene de una radiación electromagnética, puede ser absorbida únicamente mediante cuantos discretos denominados fotones. Gracias a los trabajos anteriores de Max Planck se sabía que la energía de un fotón varía inversamente con la longitud de onda de la radiación. Por ello, un fotón con longitud de onda superior a una crítica carece de la energía suficiente como para ser capaz de arrancar del metal un electrón.

Un intensificador de imágenes aprovecha el efecto fotoeléctrico para convertir el haz de fotones de una radiación electromagnética en un haz de electrones. Cada fotón puede estimular la emisión de un electrón, a lo sumo. Diríase, pues, que no se gana mucho con esta estrategia. Pero lo que hace posible la amplificación de la imagen es la carga eléctrica del electrón. Si se aplica un campo eléctrico, los electrones se aceleran y, con ello, su energía cinética aumenta. De esta forma, la energía total en juego que puede aplicarse a la formación de una imagen es susceptible de multiplicarse por un factor muy alto.

En su forma más sencilla, un intensificador de imagen es una válvula de vacío con dos electrodos conectados a una



FOTOGRAFIA DE UNA ESCENA NOCTURNA iluminada por luz estelar, hecha con una cámara de 35 milímetros, dotada con un intensificador de imagen de microcanales. El brillo de la fotografía corresponde aproximadamente al brillo de la escena tal como le aparecería al ojo humano a través del intensificador. El color verde del cielo y de las copas de los árboles es el de la luz fluorescente emitida por la pantalla sobre la que puede contemplarse la imagen amplificada. El color corresponde a la longitud de onda de la luz a la que el ojo es más sensible: en torno a los 540 nanómetros. La fotografía se realizó con una película para luz diurna en color, con una sensibilidad ASA de 400, exposición de un cuarto de segundo y una apertura de lente de *f*/2.8. Sin intensificador de imagen una fotografía de este tipo hubiera necesitado de cuatro a ocho minutos de exposición. El intensificador lo proporcionó Paul Lighty, de la División de Productos Electro-ópticos de ITT. El intensificador de imágenes se aplica, sobre todo, en operaciones militares nocturnas.

fuente exterior de alimentación. En este dispositivo no aparece incorporada ninguna placa de microcanales. Se enfoca una imagen sobre el fotocátodo y cada área del mismo emite electrones en un número determinado por el brillo de la imagen en dicha área. Cada electrón se acelera mediante un campo eléctrico, generado por un potencial de polarización de varios miles de volt y aplicado entre los dos electrodos. Los electrones acelerados viajan a través de la zona vacía del tubo hacia el ánodo, siguiendo unas trayectorias que son aproximadamente líneas rectas. De esta forma el flujo de electrones que alcanza el ánodo corresponde a las variaciones de brillo de la imagen en el fotocátodo. La resolución espacial de la imagen se conserva bien, ya que el fotocátodo y el ánodo están muy cercanos.

Un tubo de imagen tiene una salida visible porque su ánodo está construido con un material fluorescente. Cuando un electrón choca con el ánodo, excita e ioniza átomos en el material. Se libera entonces parte de la energía en forma de fotones de luz visible según vuelven los electrones de los átomos excitados a sus niveles usuales de energía. Los fotones emitidos constituyen la imagen que es visible sobre el ánodo fluorescente.

Como apunté anteriormente, cada fotón que choca con el cátodo puede generar como máximo un electrón; de hecho, la eficiencia del fotocátodo es aproximadamente de un 10 por ciento; es decir, sólo alrededor de un fotón, de cada 10, estimula la emisión de un electrón. Más aún, en el ánodo sólo alrededor de un 30 por ciento de la energía suministrada por los electrones acelerados se convierte en luz. A pesar de ese escaso rendimiento hay una ganancia neta, o amplificación de la luminosidad de la imagen, ya que se incrementa en varios miles la energía de cada electrón emitido.

Cabría imaginar que una señal imagen puede amplificarse a cualquier cuantía deseada incrementando el potencial de polarización que acelera los electrones. Desgraciadamente, un tubo de imagen con dos electrodos no puede trabajar más allá de un cierto potencial o voltaje máximo aplicado. Cuando el voltaje excede este nivel, la carga comienza a fluir espontáneamente a través del espacio entre el fotocátodo y el ánodo y enmascara la corriente más pequeña de electrones que transmiten la imagen. Puede obviarse ese efecto rediseñando el tubo de suerte que se abra un mayor espacio entre los dos electrodos; pero entonces se deterioraría la resolución espacial, debido a pequeñas desviaciones aleatorias en la trayectoria seguida por los electrones. Si se enfocan los electrones con campos eléctricos y magnéticos, podremos ensanchar la distancia entre electrodos. Se trabajará así con un voltaje más alto y se incrementará, por tanto, la ganancia en el brillo de la imagen. Los campos eléctricos y magnéticos deben, sin embargo, controlarse cuidadosamente; tales dispositivos de enfoque pueden causar distorsiones geométricas de la imagen o producir un brillo no uniforme. No obstante, para muchas aplicaciones en física y astronomía, hasta los intensificadores más complejos que enfocan los electrones tienen una ganancia insuficiente.

El hecho de que cada fotón incidente induzca la transferencia de un electrón como máximo al ánodo constituye una limitación fundamental en tales tubos de imagen. Pero hay ciertos materiales donde los electrones pueden ser arrancados tanto por un electrón entrante como por un fotón incidente. A este efecto se le llama emisión superficial secundaria y puede aprovecharse en un tubo de imagen para multiplicar el número de electrones que alcanzan el ánodo.

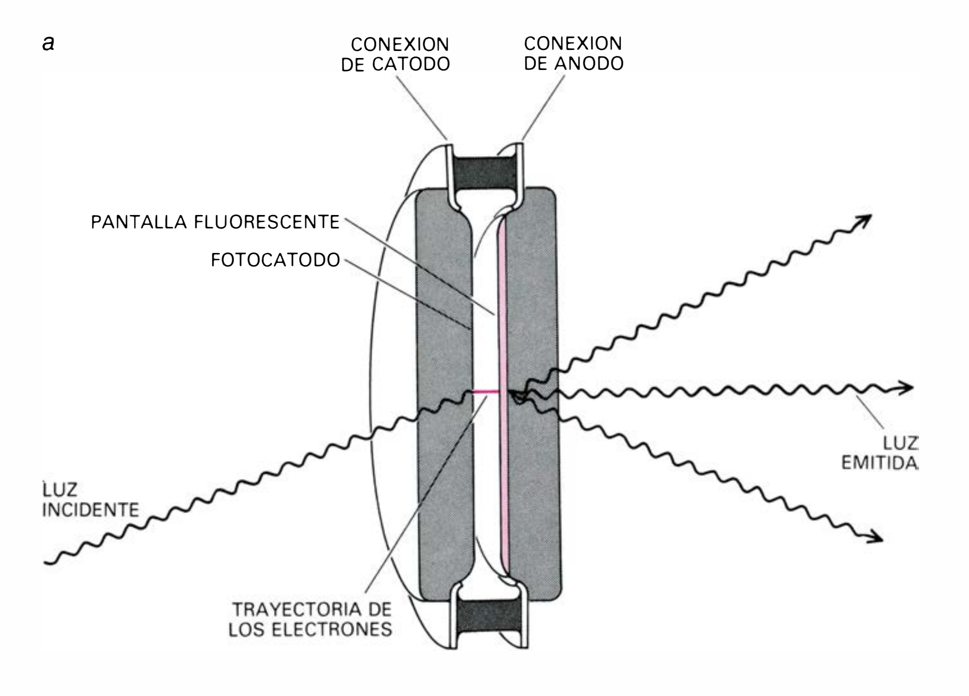
Desde los primeros años de la década de 1950 se han venido empleando multiplicadores electrónicos como componentes de amplificación en tubos fotomultiplicadores. A los electrodos metálicos, llamados dinodos, se les aplican potenciales positivos progresivamente más altos; están distribuidos de tal forma que los electrones que abandonan un dinodo son dirigidos hacia el siguiente. Un electrón incidente que choca con el primer dinodo arranca del metal varios electrones de baja energía. Se aceleran los electrones secundarios hacia el segundo dinodo, donde de nuevo se multiplican. Sucesivas multiplicaciones nos llevan a un crecimiento exponencial en la carga eléctrica liberada por el electrón inicial. Por acelerar el campo todos los electrones, la energía combinada de los electrones que finalmente chocan con el ánodo puede ser muchísimo mayor que la de un único electrón acelerado.

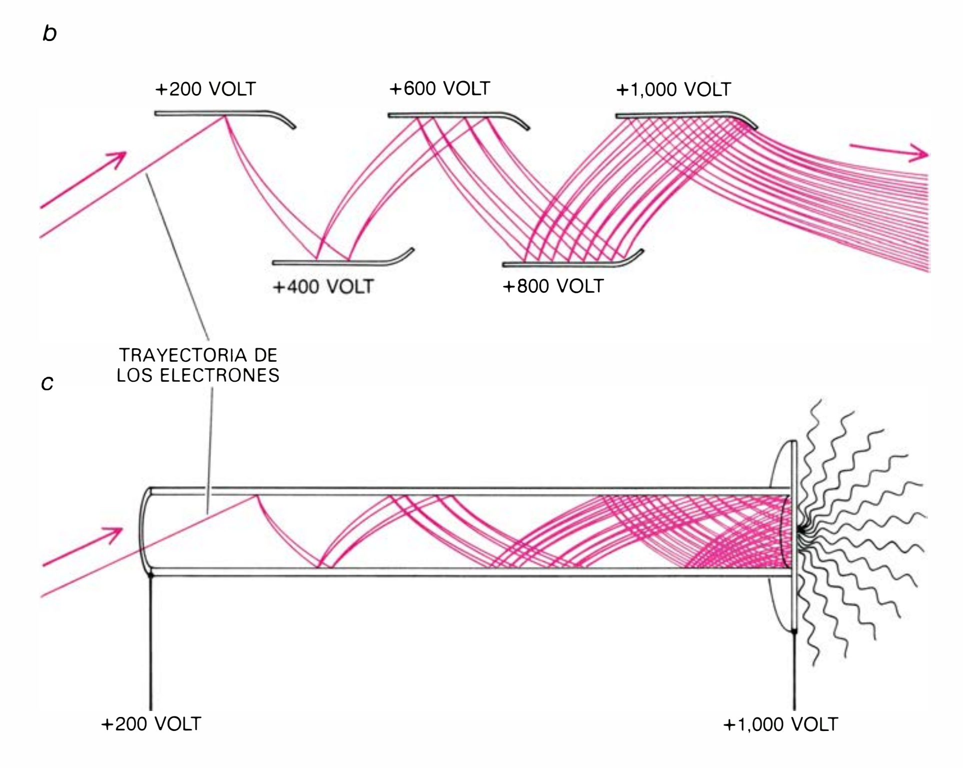
Se puede obtener una alta amplificación en una estructura multiplicadora con dinodos, pero sólo a expensas de la resolución espacial. La geometría de un dinodo no asegura que los electrones viajen en línea recta; de ello resulta que los electrones emitidos desde una región dada del fotocátodo no aterrizan necesariamente en la región correspondiente del ánodo. La placa de microcanales es un dispositivo que combina la ganancia de un multiplicador electrónico con la resolución espacial de un intensificador de imagen. Consta de millones de multiplicadores electrónicos independientes, dispuestos sobre una estructura bidimensional. No se pretende mantener la resolución espacial en cada uno de los multiplicadores; antes bien, cada uno de ellos corresponde a un único pixel o unidad de imagen. La imagen intensificada constituye una distribución de puntos diminutos de brillo variable.

La etapa clave en el desarrollo de la placa de microcanales estuvo en la invención del multiplicador electrónico de canal continuo. Se trata de un tubo de vidrio, en el que se ha hecho el vacío, conectado en cada extremo a una fuente de alimentación, que proporciona un potencial de polarización de aproximadamente 1000 volt. Actúa en forma análoga a como lo hacía el multiplicador con dinodos. Un electrón incidente inicia una emisión en cascada de electrones secundarios cuyo número crece exponencialmente según avanzan por el interior del tubo. Esta cascada de electrones está completamente confinada al interior del tubo o canal. Tales multiplicadores de canal continuo los investigaron, por primera vez, George W. Goodrich, James R. Ignatowski y William C. Wiley, de los Laboratorios de Investigación Bendix en Southfield, Michigan, de 1959 a 1961.

Los materiales destinados a la fabricación de multiplicadores de canal han de satisfacer dos requisitos. Primero, la pared del canal debe ser capaz de emitir más electrones de los que absorbe. Diversos materiales, entre los que se incluyen muchos vidrios, emiten una media de alrededor de dos electrones por cada electrón incidente, dentro del margen de energías electrónicas de fácil obtención. Segundo, la conductividad eléctrica del material tiene que ser predecible y controlable, de suerte que la carga que abandona la pared del canal pueda restituirse y restablecerse, por tanto, el campo eléctrico uniforme. Este segundo requisito significó un reto considerable para los primeros investigadores del tema.

En un vidrio, la corriente eléctrica puede conducirse por difusión de electrones libres o por la de iones (átomos cargados). Para obtener la conducción controlada, debe minimizarse la conductividad iónica del vidrio. La difusión de iones es un tipo de electrolisis





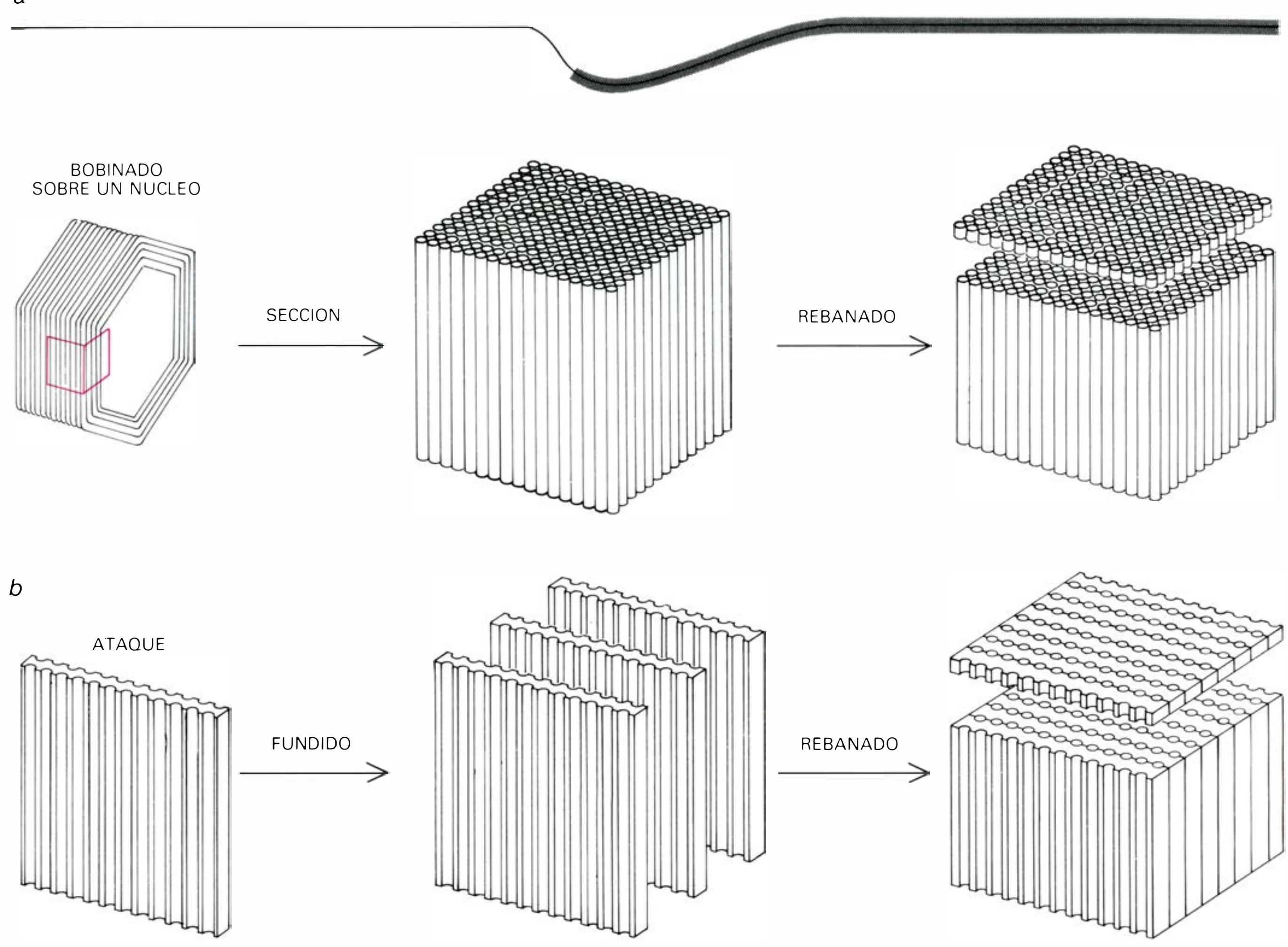
INTENSIFICADORES DE IMAGEN. Se fundan en el aprovechamiento del efecto fotoeléctrico, por el que la energía de los fotones incidentes, o cuantos de radiación electromagnética, se convierte en la energía de los electrones en movimiento. Ya que los electrones, en contraposición a los fotones, son partículas cargadas eléctricamente, su energía puede incrementarse por aceleración en un campo eléctrico. Cuando los fotones chocan con una placa metálica fotosensible, los electrones son arrojados a la región de vacío del tubo. Los electrones son acelerados hacia la placa fluorescente, cargada positivamente, donde su energía se convierte de nuevo en energía de luz visible. En un tubo de imagen de dos electrodos (*a*), se acelera, a lo sumo, un electrón a través de la región vacía por cada fotón incidente; el electrón debe viajar directamente a lo largo de toda la región vacía para que no se deteriore la imagen. Por tanto, para mantener la calidad de la imagen, se tratará de una región estrecha. La ganancia obtenible viene limitada por el voltaje aplicable entre los extremos de la zona vacía, sin que llegue a haber descarga. Los electrones que chocan con algunos materiales pueden producir electrones adicionales: se puede desencadenar una avalancha de electrones utilizando varios electrodos a los que se aplican voltajes cada vez más altos (*b*). Aunque un multiplicador electrónico de este tipo llega a rendir una ganancia alta, la dispersión de los electrones en avalancha destruye la información de la distribución espacial de los fotones incidentes. Para obtener alta ganancia y resolución espacial se recurre a un multiplicador electrónico para cada pixel, o elemento de imagen, en una figura, si bien entonces se exige que se construyan los multiplicadores muy pequeños. Esto se puede realizar con un multiplicador electrónico de canal continuo (*c*), que funciona en forma análoga al de múltiples electrodos, aunque confina la avalancha electrónica al hueco del canal. Hay varias técnicas para trefilar fibras de vidrio en la realización de multiplicadores que sólo tengan unos micrometros de diámetro; se instalan luego estructuras con multiplicadores en los tubos de imagen.

que resulta en un fenómeno químico no deseado en los electrodos, piénsese en la degradación de los metales alcalinos, y la polarización de cargas eléctricas dentro del vidrio. Por tanto, se ha de fomentar únicamente la conducción de electricidad por movimiento de electrones libres.

Aunque se han empleado diferentes clases de vidrio en la fabricación de multiplicadores de canal, la que ha tenido más éxito ha sido una mezcla de aproximadamente un 50 por ciento de óxido de plomo, un 40 por ciento de dióxido de silicio y pequeñas cantidades de varios óxidos alcalinos. En ausencia de un tratamiento más elaborado, el vidrio fabricado con la composición anterior ofrece una resistividad eléctrica muy alta o, en otras palabras, una baja conductividad. Los átomos de plomo en el vidrio son estables, bien como átomos neutros o bien como iones positivos que han cedido electrones al formar un enlace químico iónico con el oxígeno. La conductividad puede mejorarse separando del plomo el oxígeno en el tipo de reacción química denominada reducción.

La reducción puede llevarse a cabo en la superficie del vidrio calentándolo en una atmósfera de gas de hidrógeno. Después de varias horas a 400 grados Celsius, la reducción del óxido de plomo ha penetrado a una profundidad de varias décimas de micra y la mayor parte del plomo reducido se ha evaporado de la superficie. El plomo que permanece en dicha superficie se agrupa en formaciones arracimadas metálicas, dando al vidrio su característico color negro. Los electrones que habían estado formando el enlace iónico en el óxido de plomo, se hallan ahora libres para transportar corriente eléctrica; la capa superficial del vidrio se ha convertido en un semiconductor. Su resistividad se encuentra entre 10^8 y 10^{14} ohm por cuadro, margen conveniente para la fabricación de multiplicadores de canal y placas de microcanales. (La resistividad superficial se expresa en ohm por cuadro, sin especificar las dimensiones del cuadro, porque la resistividad de un cuadrado sobre una superficie de composición uniforme no depende del tamaño de dicho cuadro.)

Las primeras estructuras experimentales de multiplicadores de canal continuo se ensamblaron uniendo entre sí unas cuantas docenas de multiplicadores de canal con un vidrio que tenía un bajo punto de fusión. Cada canal venía a medir un milímetro de diámetro; la secuencia espacial de puntos que po-



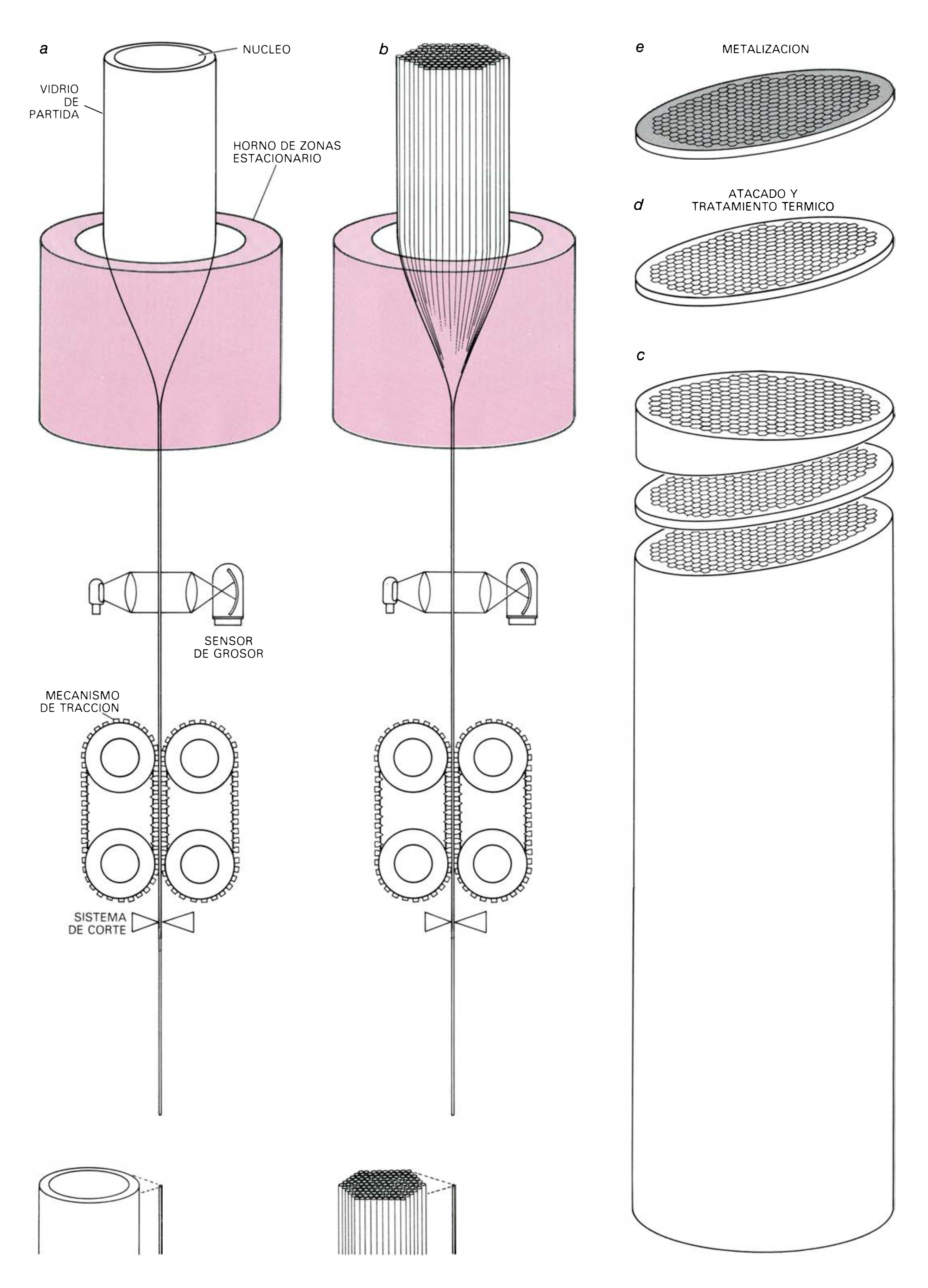
TECNICAS NO SATISFACTORIAS propuestas para fabricar tubos finos de vidrio y para su disposición en grandes estructuras, antes de que se desarrollara un método práctico. En una de esas técnicas (*a*), se recubría un alambre con vidrio fundido; enfriada la fibra resultante, se arrollaba sobre un carrete. Se cortaba una sección de las fibras densamente empaquetadas sobre el carrete y la sección se rebanaba en obleas. Finalmente, se disolvían con ácido los núcleos de alambre, dando lugar a una oblea de vidrio perforado. Resultaba difícil, sin embargo, mantener la uniformidad de las fibras durante el enrollado, y la necesidad de manipular fibras microscópicas desde la primera etapa del proceso le restaba a éste rentabilidad económica. Una segunda técnica (*b*) empleaba fotolitografía para atacar placas finas de vidrio, que eran después apiladas y fundidas para formar el bloque, a partir del cual se cortaran, a rebanadas, las obleas. Mas ni la profundidad ni la anchura de los surcos podía controlarse en el proceso de ataque dentro de las tolerancias debidas.

día diferenciarse con estos primitivos instrumentos resultaba, pues, bastante pequeña. Por otra parte, se necesitaban un gran número de elementos amplificadores para reproducir una imagen de alta calidad. Se construyó una estructura, laboriosamente ensamblada a mano, en los Laboratorios de Investigación de la Bendix; constaba de unos 5000 canales, empaquetados hexagonalmente, con una separación de 0,15 milímetros entre los centros de cada canal. Aunque la resolución de puntos con dicha estructura excedía la calidad usualmente alcanzada por periódicos y revistas en la reproducción de fotografías, el área de la imagen rondaba sólo en torno al centímetro cuadrado.

Podría parecer que no cabe extrapolar el resultado de experimentos con dispositivos dentro de la escala de los primitivos instrumentos a las estructuras microscópicas de los multiplicadores de canal, capaces de transmitir imágenes de alta calidad. Pero la extrapolación es posible en parte, ya que, para un potencial aplicado dado, la ganancia de un multiplicador de canal no depende de su tamaño. En un canal más largo, cada electrón llega más lejos bajo la influencia de la fuerza ejercida por el campo eléctrico y, por tanto, tiene más tiempo para ganar energía. En un canal más corto el campo eléctrico es más intenso: la fuerza que actúa sobre el electrón varía inversamente con la distancia entre los electrodos. Los dos efectos se cancelan entre sí.

Las estructuras fabricadas a mano establecieron el hecho de que los canales individuales en una estructura responden independientemente entre sí. Se encontró, sin embargo, que la calidad de la imagen es bastante sensible a las irregularidades en su espaciamiento. Al principio de los años 1960, el principal reto al desarrollo de las placas de microcanales era el de idear técnicas de fabricación que pudieran mantener una gran uniformidad en el diámetro de los canales.

En los Laboratorios de Investigación de la Mullard en Londres se inventó la técnica del proceso de núcleo metálico, que desarrollaría luego el Laboratorio de Electrónica y Física Aplicada de París. Se recubre un alambre uniforme y fino con vidrio multiplicador de canal ablandado por calentamiento; se arrolla a un tambor poligonal. Una vez enrollada una capa suficientemente gruesa, se corta en bloques. Estos bloques se apilan y se funden en un bloque de vidrio multiplicador, que contiene así millares de alambres. Se corta este bloque, finalmente, en rodajas muy finas y se ataca con ácido diluido para eliminar los alambres metálicos. Las obleas de vidrio resultantes están horadadas por



a
NUCLEO
VIDRIO
DE
PARTIDA
HORNO DE ZONAS
ESTACIONARIO
b
e
METALIZACION
d
ATACADO Y
TRATAMIENTO TERMICO
c
SENSOR
DE GROSOR
MECANISMO
DE TRACCION
SISTEMA
DE CORTE

millares de canales de diámetro muy uniforme.

El proceso de núcleo metálico presenta dos desventajas importantes. Primera, aunque el diámetro de los canales sea uniforme, no lo es su espaciamiento. Es difícil arrollar el frágil alambre revestido de vidrio alrededor del núcleo con perfecta uniformidad. La segunda desventaja es que, en todas las etapas de fabricación de la fibra de vidrio, ésta tiene su tamaño final microscópico. Por tanto, las máquinas de revestimiento y de arrollamiento, que representan una inversión de capital considerable, pueden procesar sólo una pequeña cantidad de material por unidad de tiempo, lo que constituye una grave objeción económica.

Otra técnica, llamada proceso de placa estriada, fue objeto de investigación en los Estados Unidos y en Suiza. En placas de vidrio delgadas, se grava, por fotolitografía, un gran número de finas estrías paralelas. Se apilan las placas y se funden para formar el bloque a partir del cual se corten las obleas de microcanales. Pueden fabricarse además canales curvados o en zigzag, ya que estas configuraciones superan ciertas limitaciones de ganancia que son inherentes a la geometría de canales rectilíneos. Sin embargo, el difícil control de la anchura y la profundidad de las estrías durante el atacado y la fusión, así como las desigualdades resultantes en la ganancia, llevaron a un abandono eventual del proceso de placa estriada.

La técnica que se ha demostrado más rentable, desde el punto de vista económico, para la fabricación en masa de placas microcanales, es la del proceso de doble trefilado. Se basa en los métodos usados para estirar el vidrio en finas fibras microscópicas; procedimientos que conocían ya los talleres egipcios de la XVIII dinastía (1570-1320 a. de C.). Algunas composiciones de vidrio, que pueden realizarse en un rango muy amplio de temperaturas, tienen la propiedad de preservar su sección transversal en las fases de calentamiento y estirado. Los investigadores de la Bendix fueron los primeros en sacarle partido, en 1960, a esa propiedad a la hora de fabricar placas de microcanal; a partir de entonces la han venido usando otras empresas.

En el proceso de trefilado doble, el vidrio del multiplicador de canal se extrusiona o estampa, obteniéndose un lingote cilíndrico. A continuación, se enfría y pule para formar una varilla uniforme de varios centímetros de diámetro. Se horada la varilla según su eje, introduciéndose, por lo general, en el hueco resultante, otra varilla, también de vidrio, pero con una composición diferente. Cuando finaliza el proceso, se elimina este núcleo mediante atacado en un baño de ácido diluido y caliente. Durante las etapas intermedias de fabricación, dicho núcleo ha tenido la misión de soportar al cilindro externo al objeto de que los microcanales ya acabados posean tamaños y figuras prácticamente uniformes.

Mientras transcurre el primer trefilado, el cilindro de vidrio se suspende verticalmente en un horno de zonas, donde la temperatura puede controlarse punto por punto. La base del cilindro se calienta a unos 500 grados Celsius; una gota de vidrio reblandecido, suspendida por una fibra de vidrio cuyo diámetro es aproximadamente de un milímetro, desciende desde el horno. Cuando la fibra lleva ya descendidos varios metros por debajo del horno, está lo suficientemente fría como para ser manejada por una máquina de tracción que regula la velocidad del trefilado. A su vez, la regulación de la velocidad da un control preciso del diámetro de la fibra. Por debajo de la máquina de tracción, la fibra se corta en segmentos de unos 15 centímetros de largo. Varios miles de estos segmentos se ensamblan en un mazo o haz hexagonal.

El segundo trefilado es similar al primero. El paquete hexagonal de fibras se suspende y se calienta en un horno de zonas, trefilándose en una fibra compuesta hexagonal, de un milímetro de espesor. Así pues, el espaciamiento entre los distintos cilindros queda reducido a unas pocas centésimas de milímetro, siendo éste, en consecuencia, el espaciamiento final en la placa de microcanal. De nuevo, el vidrio trefilado se corta en segmentos que son empaquetados juntos y fundidos en el vacío para formar un bloque de hexágonos, bloque que puede medir hasta 125 milímetros de diámetro e incorporar millones de microcanales.

Rebanando el bloque en obleas de un grueso de un milímetro, y puliendo las caras de cada oblea, se obtienen placas microcanales. El corte se hace normalmente en ángulo oblicuo a los ejes de los microcanales, de suerte que, en la placa acabada, colisionen los electrones con la pared del canal cerca del extremo de entrada, en vez de ir directamente al ánodo. Procesos alternativos como trituración, taladrado y templado pueden preparar la oblea para que satisfaga necesidades especiales. En las etapas finales, el vidrio del núcleo se disuelve en ácido y la placa se calienta en gas de hidrógeno para reducir el óxido de plomo, dando a la superficie del vidrio la conductividad eléctrica requerida. La oblea se remata evaporando una película metálica en ambas caras, de modo que puedan operarse las conexiones eléctricas en todos los canales.

Existen dos tipos básicos de intensificadores de imagen que incorporan una placa microcanal. El intensificador enfocado por proximidad se asemeja a un intensificador simple de dos electrodos en el que se ha insertado una placa microcanal. Se enfoca la imagen incidente en el fotocátodo, los electrones emitidos se aceleran a través de una pequeña abertura en la placa microcanal y la mayor parte de los electrones que emergen de la placa se aceleran, a través de una abertura algo mayor, hasta una pantalla fluorescente. Una válvula de este tipo es compacta, libre de distorsión y no se ve afectada por campos magnéticos. Más aún, la imagen de salida es plana y está siempre enfocada del mismo modo que la imagen formada por impresión sobre papel fotográfico.

El intensificador enfocado por proximidad vale, sobre todo, para realizar fotografías de fenómenos de alta velocidad. La placa de microcanales intensifica la imagen tenue que resulta de tiempos de exposición fotográfica breves. Además, el dispositivo puede servir como una cámara de obturador rápido: la diferencia de voltaje, controlada externamente, entre el fotocátodo

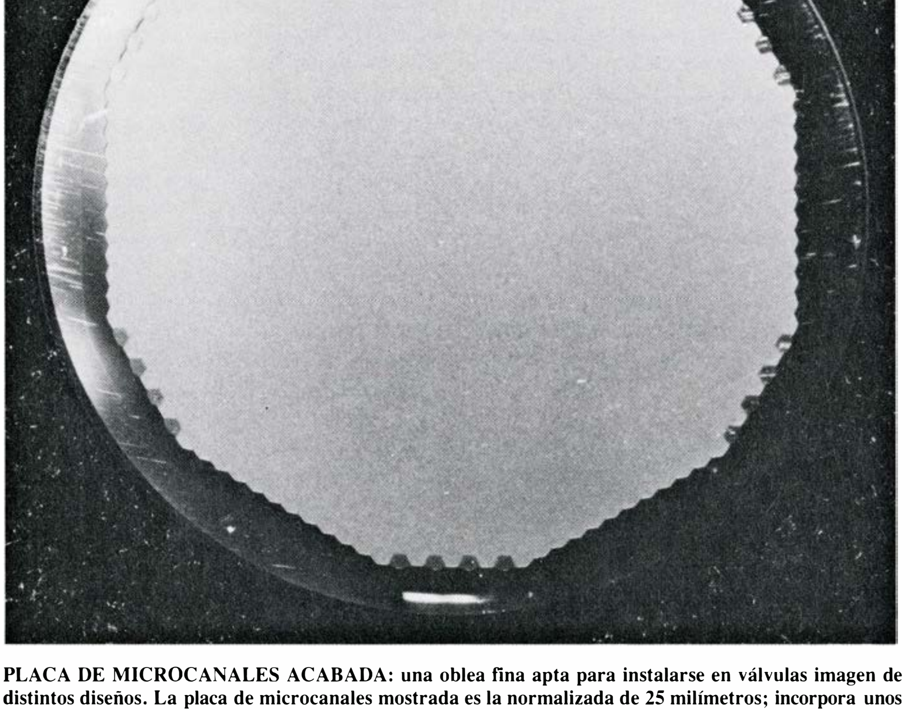
PROCESO DE DOBLE TREFILADO para generar placas de microcanales. Se trata del método de trefilado de fibras a partir de un bloque de vidrio calentado. Las fibras mantienen la geometría de la sección transversal del bloque, aunque en menor escala. Se rellena un tubo cilíndrico de vidrio, de unos 50 milímetros de diámetro, con un núcleo también de vidrio que puede eliminarse luego por ataques con ácidos. El tubo se calienta y trefila hasta un grosor de alrededor de un milímetro (*a*). Varios miles de tubos trefilados, de unos 15 centímetros de longitud, se ensamblan manualmente en un manojo hexagonal de 50 milímetros. El bloque que resulta se calienta de nuevo y trefila hasta un grosor de un milímetro aproximado (*b*). La fibra resultante constituye una réplica geométricamente similar al hexágono inicial de 50 milímetros, formando por el conjunto de tubos ensamblados de vidrio. Las fibras hexagonales se unen, a continuación, por fusión; se cortan con un ángulo oblicuo a fin de formar obleas de alrededor de un milímetro de espesor (*c*). Estas obleas, atacadas en un baño de ácido que elimina los núcleos de vidrio, se calientan después en una atmósfera de hidrógeno que genere la conductividad eléctrica deseada en las paredes internas de los microcranales (*d*). Se evapora luego un metal sobre ambas caras de la oblea, con lo que pueden efectuarse las conexiones eléctricas (*e*). Los tamaños de los tubos empaquetados en el hexágono y los hexágonos ensamblados en la oblea concluida se han exagerado aquí para una mayor claridad.

y la parte delantera de la placa, se aprovecha para conmutar rápidamente el intensificador. Un intensificador enfocado por proximidad puede responder a pulsos eléctricos de disparo cuya duración es de unos pocos nanosegundos. Por tanto, puede fotografiarse nítidamente un acontecimiento de muy pequeña duración.

La segunda clase de intensificador de imagen emplea un campo electrostático no uniforme para enfocar los electrones que emite el fotocátodo hacia la placa de microcanales. Como en las lentes ópticas, el tubo invierte la imagen y puede también ampliarla; en algunos casos el dispositivo actúa como una lente zoom, es decir, se ajusta simultáneamente la amplificación y el foco. Debido a la inversión de la imagen, el intensificador enfocado electrostáticamente es útil, en particular, cuando se usa junto con dispositivos ópticos que también invierten la imagen. La placa de microcanales puede reemplazar un sistema óptico complejo, así el empleado en los binoculares, restaurando a su posición original la imagen invertida que resulta del uso de una lente ordinaria.

Aunque la placa de microcanales se ha aplicado fundamentalmente a la intensificación de imágenes, puede amplificar cualquier forma de radiación capaz de iniciar una avalancha de electrones en los canales. Los iones y los electrones, de igual forma que los fotones de longitudes de onda en las regiones del espectro del ultravioleta y rayos X, pueden desencadenar avalanchas en los microcanales. Cabe, pues, percibir esas formas de radiación sin utilizar un fotocátodo. Se enfoca la radiación sobre la cara delantera de la placa y los electrones generados se detectan en la cara posterior. Si la placa está equipada con una pantalla fluorescente, que haga de ánodo, puede funcionar como un convertidor de imagen, haciendo visible la secuencia espacial de radiación que se encontraba fuera de la región visible del espectro.

Las placas de microcanales, que actúan tanto como convertidores como intensificadores, han sido incorporadas a instrumentos de laboratorio: osciloscopios de alta velocidad, microscopios electrónicos de transmisión y espectrómetros de masa. En estas aplicaciones, la ganancia de la placa de microcanales posibilita el registro de características de bajo contraste o débiles y extremadamente breves. Más aún, permite que los instrumentos trabajen con una corriente de haz más baja, lo que representa a veces una ventaja importante. En un microscopio electrónico, por ejemplo, se puede reducir el daño causado por el bombardeo electrónico al material sujeto a examen.



PLACA DE MICROCANALES ACABADA: una oblea fina apta para instalarse en válvulas imagen de distintos diseños. La placa de microcanales mostrada es la normalizada de 25 milímetros; incorpora unos tres millones de microcanales que pueden multiplicar el brillo de una imagen por un factor de 10.000. La placa microcanal fue fabricada por la Compañía Electro-óptica Galileo en Sturbridge, Massachusetss.

Cuatro restricciones físicas importantes limitan el comportamiento de las placas de microcanales de corriente. La primera es que la señal de salida media no puede exceder la máxima corriente que pueden sustentar las paredes del microcanal. Cuando el flujo de electrones es demasiado grande, la carga eléctrica arrancada del vidrio no se sustituye inmediatamente y, por tanto, se modifica el campo; de resultas de ello, se reduce la ganancia total del canal. Las partes brillantes de una imagen pueden llegar a saturarse y se pierde el contraste en estas regiones. Aunque este efecto no es deseable en astronomía, donde se necesita un margen dinámico muy bien calibrado, sí es beneficioso en dispositivos de visión nocturna, donde un gran contraste en las partes brillantes de una imagen puede disminuir la sensibilidad del ojo del observador en las regiones más oscuras.

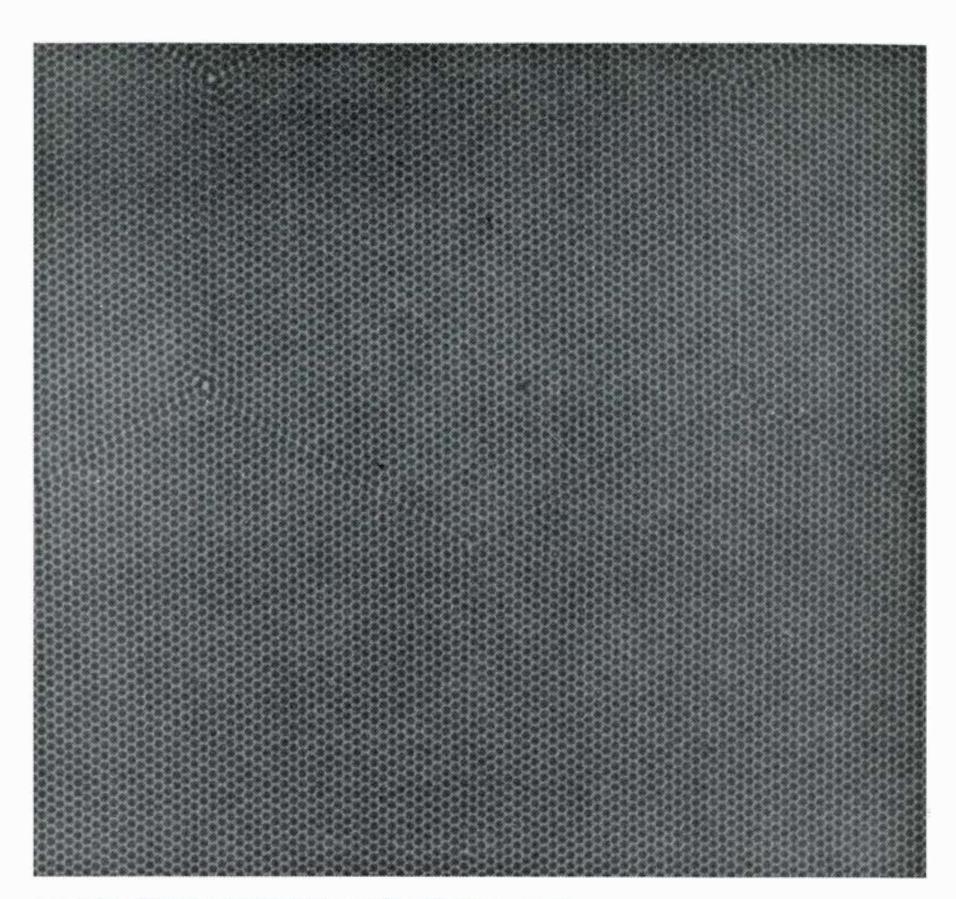
Una segunda limitación de los multiplicadores de canal la tenemos en un fenómeno llamado realimentación iónica. Cuando el canal trabaja con una alta ganancia, los átomos de gas del canal pueden ionizarse por colisiones con la avalancha de electrones. (Al menos unos pocos átomos están siempre presentes incluso en el vacío mejor conseguido.) Un ion cargado positivamente, y formado en esta forma, es acelerado por el campo eléctrico hacia el extremo de entrada del tubo, donde puede chocar con la pared del canal e iniciar una nueva avalancha. Esa nueva cascada puede desencadenar nuevas ionizaciones; los electrones adicionales tienden a enmascarar la señal. Mas aún, los iones que chocan con el fotocátodo pueden acortar su vida, aunque ahora se evitan rutinariamente estas averías depositando una fina capa de óxido de aluminio, permeable a los fotoelectrones pero no a los iones, sobre la cara de entrada a la placa.

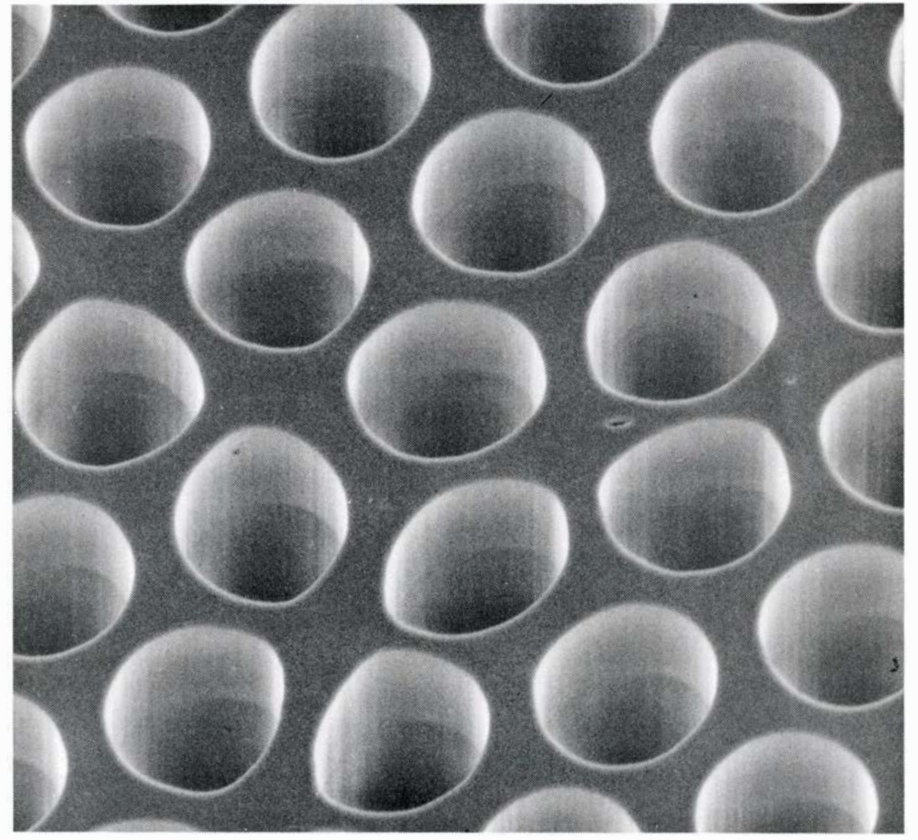
¿Cómo minimizar la realimentación iónica? Por construcción de canales curvados o en forma de zigzag. Los electrones se dirigen fácilmente ha-

cia el ánodo, pero se inhibe el movimiento de los iones hacia el fotocátodo por las colisiones con las paredes. Aunque la forma más sencilla de cortar los canales es por grabado fotolitográfico, el proceso, como ya se ha indicado, no resulta satisfactorio desde un punto de vista técnico. En las aplicaciones donde se necesita alta ganancia se ponen a menudo juntas dos o más placas y se giran una respecto a la otra. Ya que las caras de la oblea no están cortadas en ángulo recto con respecto al núcleo hueco del canal, cada microcanal resultante se halla parcialmente curvado. Las sucesivas avalanchas de electrones siguen el camino curvado, no así los iones, al no desencadenar avalanchas de otros iones. Tal estrategia permite incrementar sustancialmente el voltaje aplicado (y por tanto la ganancia) antes de que se alcance el umbral de realimentación regenerativa de iones.

En un esfuerzo por realizar microcanales curvados, y minimizar así la realimentación iónica, en varios laboratorios están utilizando ya la variación controlable de la viscosidad del vidrio con la temperatura. Si una placa microcanal de vidrio se empareda entre dos bloques de metal que están a diferentes temperaturas, se establece un gradiente uniforme de temperatura en la placa. El vidrio cercano al bloque más caliente estará más fluido. Si la placa se "plancha" por desplazamiento lateral de uno de los bloques, los microcanales están más ladeados cerca del bloque más caliente que los que lo están cerca del bloque más frío. Por tanto, cada microcanal aparecerá curvado. Aunque las placas de canales curvados son todavía dispositivos experimentales, pueden multiplicar los electrones secundarios por un factor de hasta un millón sin realimentación iónica, y presentan la misma resolución de rasgos espaciales que las placas de microcanales convencionales.

Una tercera limitación de un multiplicador de canal es la densidad de carga de los electrones en el canal: la denominada carga espacial. Cuando la avalancha de electrones alcanza una densidad lineal de unos 10 millones de electrones por milímetro, la repulsión electrostática mutua tiende a rechazar los electrones secundarios adicionales hacia la superficie del canal antes de que el campo pueda acelerarlos significativamente. El resultado es la limitación del número de electrones en la avalancha. En muchas aplicaciones este efecto resulta beneficioso, ya que hace casi idénticos los pulsos de salida del



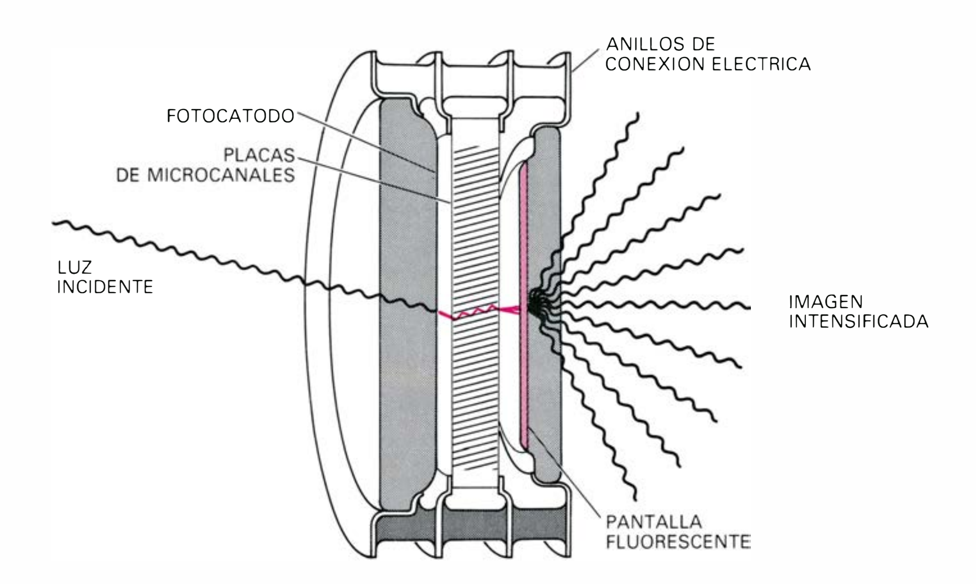


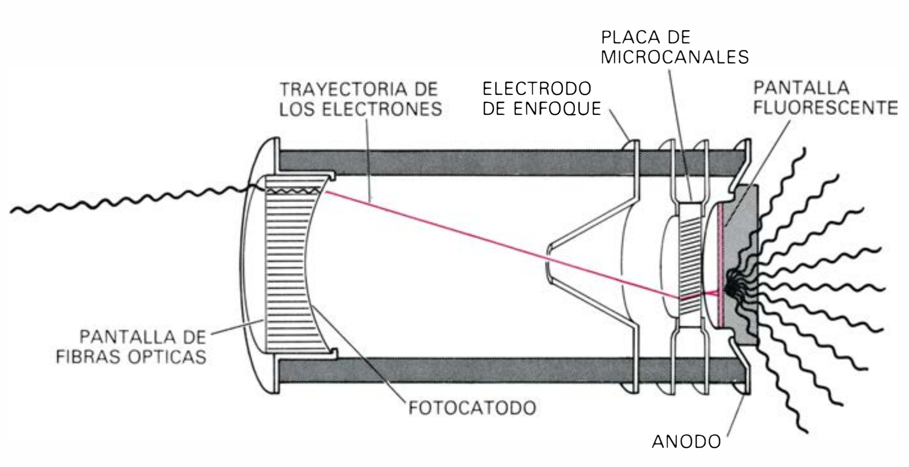
UNIFORMIDAD DE TAMAÑO Y DE ESPACIAMIENTO de los microcanales, propiedades críticas en la transmisión de imágenes de alta calidad. Una placa microcanal hecha de fibras de vidrio trefilado presenta una buena uniformidad, como se muestra en estas dos fotomicrografías. En la superior, hecha por Adolf R. Asam de la División de Productos Electro-ópticos ITT, vemos una placa de microcanales aumentada 110 diámetros; su estructura secundaria hexagonal se distingue con absoluta nitidez. En la fotomicrografía inferior, hecha por el autor, una segunda placa de microcanales está aumentada 650 diámetros. Estos microcanales miden 40 micras de diámetro y se hallan espaciados entre sí otros 50 micrometros.

multiplicador. Se suprimen con ello las fluctuaciones aleatorias en la carga total de la avalancha de electrones.

La cuarta limitación de la placa microcanal está relacionada con su rendimiento. El área de la superficie de las entradas de los canales es inferior al área total de la superficie de la placa. Si se emplean canales circulares, la relación de la sección transversal al área de la placa debe ser menor que un 91 por ciento, por restricciones de tipo geométrico; debido al grosor de las paredes del canal, la relación en la mayoría de las placas microcanal se cifra sólo en un 55 por ciento. Por tanto, la mitad aproximada del flujo de entrada (input) choca con el área metálica entre las entradas de los canales.

Se pueden adoptar ciertas estrategias para reducir esta pérdida. Si se aplica un campo electrostático fuerte a la cara frontal de la placa, los electrones emitidos desde las áreas de tela metálica pueden arrastrarse hasta un canal adyacente e iniciar así una avalancha. Las entradas de los canales pueden construirse también en forma de embudo mediante atacado de la oblea, incrementando así la relación entre el área de los canales y el área total. Varios laboratorios están experimentando con canales hexagonales o cuadrados, que se empaquetan de una forma más eficaz que los de sección transversal circular.





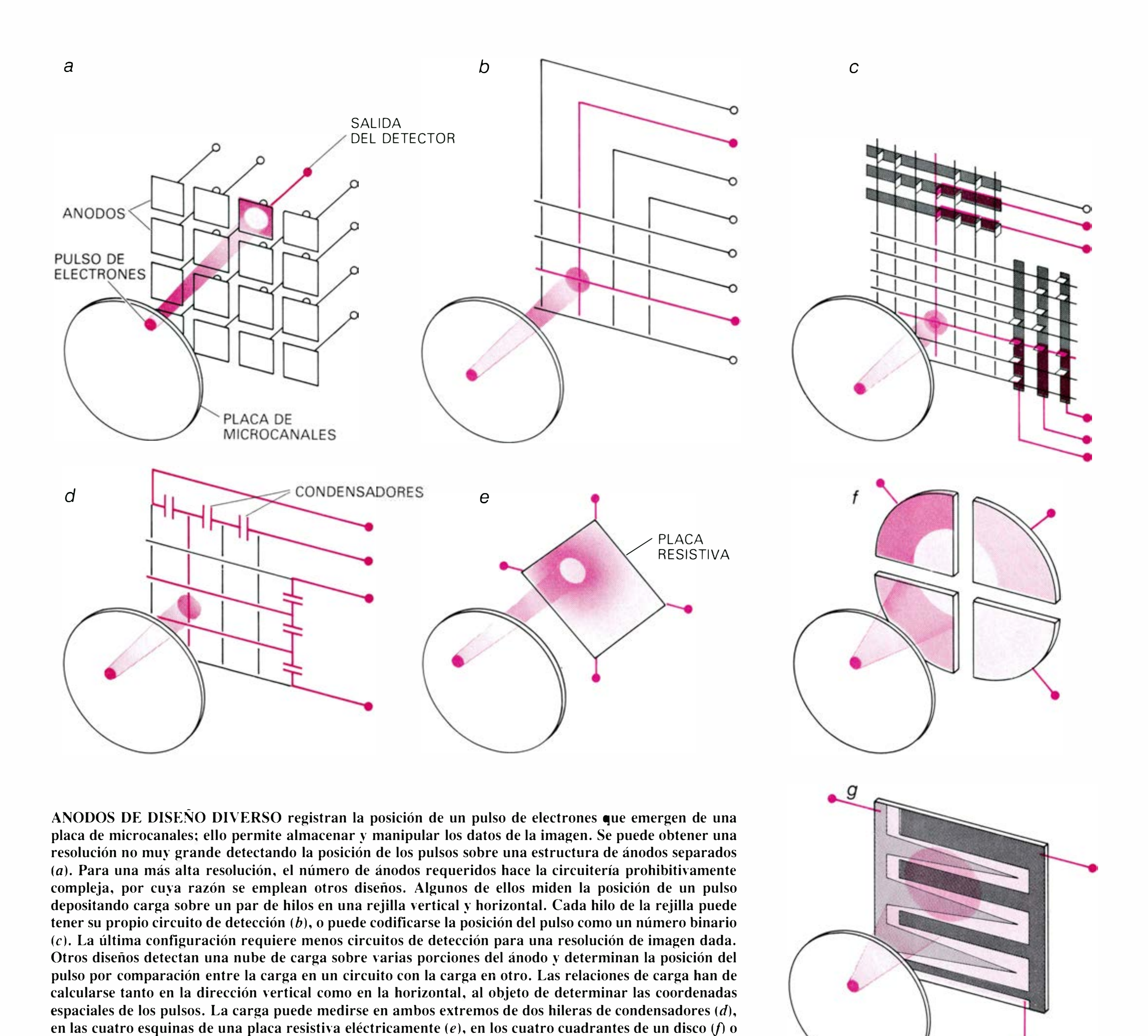
TUBOS DE IMAGEN, que incorporan las placas de microcanales en dos configuraciones principales. En cada una de ellas, la placa queda emparedada entre un fotocátodo o electrodo negativamente cargado y un ánodo fluorescente o electrodo cargado positivamente. Los potenciales eléctricos de los electrodos sobre la placa de microcanales son intermedios entre el del fotocátodo y el ánodo; así pues, los potenciales de los cuatro electrodos del diagrama van siendo más positivos de izquierda a derecha. El intensificador de enfoque cercano (*diagrama superior*) es similar al intensificador de dos electrodos. Los electrones emitidos por el fotocátodo cruzan el corto espacio vacío hasta la placa de microcanales, donde son acelerados y multiplicados. Los electrones que emergen desde la placa vuelven a chocar con el ánodo a través del espacio vacío. La distribución espacial de la imagen de entrada se conserva merced a que las dos zonas vacías son suficientemente pequeñas para minimizar la dispersión electrónica. Los fotones entran en un intensificador enfocado electrostáticamente (*diagrama inferior*) a través de una gruesa ventana de fibra óptica que cierra el vacío interior del tubo o válvula imagen. En la parte interior de la ventana se halla el fotocátodo; los electrones que emite son enfocados por el campo eléctrico en la placa de microcanales. De esta forma la imagen brillante es invertida, y el tubo de imagen puede acoplarse a sistemas ópticos adicionales que también invierten las imágenes. La imagen final aparecerá, pues, en su posición correcta. Cambiando el foco del haz electrónico podremos entonces ampliar o reducir el tamaño de la imagen.

La salida de una placa microcanal no hay que registrarla sobre una pantalla fluorescente. En muchas aplicaciones, así la transmisión remota de datos fotométricos desde un ingenio espacial, se necesita una salida de tipo eléctrico en vez de una salida de tipo visual. El sensor más simple de este tipo es el fotomultiplicador de microcanales; se asemeja al intensificador microcanal enfocado por proximidad en el cual un simple ánodo metálico reemplaza a la pantalla fluorescente. Se obtiene un pulso de salida eléctrico desde el ánodo cada vez que la avalancha de electrones atraviesa la placa de microcanales. Merced al brevísimo retraso entre el estímulo y el pulso de salida amplificado, el fotomultiplicador de microcanales se ha empleado para contar fotones uno a uno y medir el tiempo de vuelo de las partículas subatómicas producidas en los aceleradores.

Los fotomultiplicadores convencionales poseen una sola conexión para su salida eléctrica; de ahí que no sirvan para la percepción de imágenes. Por otro lado, puede diseñarse un fotomultiplicador de microcanales de suerte que permita decodificar eléctricamente la posición de un pulso de salida sobre el ánodo. En teoría, la forma más simple de llevarlo a cabo es conectar a cada pixel su propio ánodo y circuito de detección. Pero se trata de una técnica limitada a la salida de imágenes bastante toscas, ya que el número de ánodos y circuitos disparadores (que debe ser igual al número de pixels) no puede incrementarse indefinidamente.

Para reducir el número de circuitos de disparo, se detecta simultáneamente cada pulso sobre rejillas solapadas de hilos verticales y horizontales. Si el pulso dispara el hilo vertical y el horizontal más próximos en las rejillas, el número de circuitos requeridos variará según la raíz cuadrada del número de pixels. Disposiciones más complicadas nos pueden llevar a codificar, con un mayor rendimiento, la información posicional. En particular, si el pulso dispara cualquier combinación de hilos en las rejillas, cada posible combinación de hilos codifica alguna posición del pulso. Con una configuración de este tipo el número de circuitos varía únicamente como el logaritmo del número de pixels.

Otra clase de ánodos sensores de po-



ANODOS DE DISEÑO DIVERSO registran la posición de un pulso de electrones que emergen de una placa de microcanales; ello permite almacenar y manipular los datos de la imagen. Se puede obtener una resolución no muy grande detectando la posición de los pulsos sobre una estructura de ánodos separados (*a*). Para una más alta resolución, el número de ánodos requeridos hace la circuitería prohibitivamente compleja, por cuya razón se emplean otros diseños. Algunos de ellos miden la posición de un pulso depositando carga sobre un par de hilos en una rejilla vertical y horizontal. Cada hilo de la rejilla puede tener su propio circuito de detección (*b*), o puede codificarse la posición del pulso como un número binario (*c*). La última configuración requiere menos circuitos de detección para una resolución de imagen dada. Otros diseños detectan una nube de carga sobre varias porciones del ánodo y determinan la posición del pulso por comparación entre la carga en un circuito con la carga en otro. Las relaciones de carga han de calcularse tanto en la dirección vertical como en la horizontal, al objeto de determinar las coordenadas espaciales de los pulsos. La carga puede medirse en ambos extremos de dos hileras de condensadores (*d*), en las cuatro esquinas de una placa resistiva eléctricamente (*e*), en los cuatro cuadrantes de un disco (*f*) o en disposiciones geométricas más complejas de ánodos aislados eléctricamente (*g*). En cada caso, las relaciones de la carga detectada por los distintos circuitos nos darán la posición de un pulso sobre la placa de microcanales dentro de un margen de unas pocas milésimas del diámetro del campo de visión.

sición codifican las coordenadas del pulso electrónico en una forma análoga. Describiremos una distribución, desarrollada recientemente, que usa sólo tres circuitos de detección. El ánodo se compone de tres electrodos metálicos, colocados muy alejados de la cara de salida de la oblea de microcanales, de suerte que cada electrodo intercepta una distribución de la nube electrónica asociada con cada electrón de la avalancha. Uno de los electrodos tiene forma de zigzag y divide el plano del ánodo en dos partes [*véase la ilustración de esta página*]. Sobre una parte del zigzag, el área del electrodo colector de carga varía linealmente con la coordenada *x* y es independiente de la coordenada *y*. El área del electrodo en la otra parte del zigzag varía linealmente con la coordenada *y*, y es independiente de la coordenada *x*. Amplificando y midiendo los pulsos de carga sobre los tres electrones independientemente, y calculando las relaciones entre los pulsos recibidos por los electrodos dependientes de *x* y dependientes de *y* y la suma de los tres pulsos, se puede calcular con precisión la posición del pulso de salida. La idea de semejante diseño se le ocurrió a Hal O. Anger, de la Universidad de California en Berkeley.

Con mis colegas del Laboratorio de Ciencias Espaciales en Berkeley, hemos adoptado el esquema de Anger en dos aplicaciones de vuelos espaciales. Otros cuatro grupos pretenden ahora incorporar el detector a fotomultiplicadores que cuentan fotones de longitudes de onda visibles. Tal detector sensible de posición puede funcionar como el dispositivo de entrada en un sistema de ordenador capaz de registrar y analizar los datos. Por tanto, la placa de microcanales puede servir como intermediaria entre la radiación entrante y unos datos de salida de variedad casi ilimitada. Se confía en que sistemas de esta clase faciliten el camino a otros avances significativos en física y astronomía.

λόγου· ἀπεπροσεστέρων. Ὡς μελλ ἡμῖν ἡρμοσμένων ἐν ταύτῃ καταστάσει· ἄξιον
δὲ προσξένοις ἡμετέροις καὶ τοῖς ἡμῶν συμμάχοις εἰθισμένοις εἶναι. καὶ θεῶν
καὶ ἡρώων ἕνεκα, καὶ χαρίτων, ἃ ἀνθρώποις παρὰ ἀνθρώπων γίνονται. ἐχθροὺς
μὲν ἐργάζεσθαι ἀλλήλων ἀδικίας· εἰς δὲ φιλίαν προτρέπεσθαι. εἰ γὰρ μὴ ἐν τῇ
ἡμετέρᾳ πόλει τούτων ὄντων σωζόμεθα· οἱ κοινὸν πονοῦντες. ὧν ἡμ-
ερομένων ἀκυρώσομεν :-

Ἐκ τοῦ κατὰ τὸν Ἱπποκράτην ὅρκου
καθόσον οἷόν τε χριστιανῷ ὀμόσαι:-
Εὐλογητὸς ὁ θεὸς καὶ πατὴρ τοῦ κυρίου ἡμῶν
Ἰησοῦ Χριστοῦ· ὁ ὢν εὐλογητὸς εἰς τοὺς αἰ-
ῶνας ὅτι οὐ ψεύδομαι. οὐ μολύ-
νω τὴν τῆς ἰατρικῆς τέχνης μά-
θησιν· οὐδὲ δώσω τινὶ αἰτηθεὶς
φάρμακον θανάσιμον· οὐδὲ ὑ-
φηγήσομαι ξυμβουλίην τοιήνδε.
ὁμοίως δὲ οὐδὲ γυναιξὶ δώσω
φθόριον· ἄνωθέν τε ἢ κάτωθεν· ἀλλὰ διδάξω τὴν τέχνην ταύτην τοῖς χρῄ-
ζουσιν μανθάνειν· ἄνευ φθόνου τε καὶ ξυγγραφῆς. καὶ διαιτήμασιν ἐπ' ὠ-
φελείῃ καμνόντων κατὰ δύναμιν καὶ κρίσιν ἐμήν. καὶ ἁγνῶς καὶ ὁσίως
διατηρήσω τέχνην ἐμήν· ἐς οἰκίας ὁκόσας ἂν εἰσίω· ἢ ἐσελεύσομαι ἐπ' ὠ-
φελείῃ καμνόντων· ἐκτὸς ἐὼν πάσης ἀδικίης· ἑκουσίης τε καὶ ἀκου-
σίης· φθορίης τε. καὶ τῆς ἄλ-
λης λοιμώδεος. καὶ ἀφρο-
δισίων ἔργων· ἐλευθέρων
τε. καὶ δούλων· ἀνδρεί-
ων τε καὶ γυναικείων σωμάτων.
ὅσα δ' ἂν ἐν θεραπείῃ· ἢ ἴδω· ἢ
ἀκούσω· ἢ ἄνευ θεραπείης σὺν
κατ' αἰῶνα ἀνθρώπων· ἃ μὴ χρὴ ἐξω-
λαλεῖσθαι· σιγήσομαι· ἄρρητα
ἡγεύμενος εἶναι τὰ τοιαῦτα. ὅρκον μὲν
οὖν μοι τόνδε ἐπιτελέα ποιέοντι
καὶ μὴ ξυγχέοντι· βοηθός μοι γένοιτο
ὁ θεὸς καὶ βίου καὶ τέχνης· δοξαζο-
μένῳ παρὰ πᾶσιν ἀνθρώποις
ἐς τὸν αἰεὶ χρόνον εὐορκοῦντι
μὲν μοι. ἀλλ' ἐπιορκοῦν-
τι δὲ τἀναντί-
α τούτου
τέ
λγτ

Hipócrates

En los orígenes de la ciencia occidental se halla el tránsito de la medicina credencial a la racional en la antigua Grecia. El Corpus hipocrático *señala la vía de la experiencia contrastada con la razón*

José Alsina

La medicina griega va unida, indisolublemente, al nombre de Hipócrates, y especialmente a una serie de escritos, transmitidos por los manuscritos medievales, que constituyen el llamado *Corpus hippocraticum*, lo que no significa que todos los opúsculos en él contenidos sean obra del famoso médico. Este es un problema que abordaremos más adelante. Pero la medicina griega ha conocido una larga historia. Unos orígenes oscuros, una gestación larga y complicada preside su definitiva constitución en la segunda mitad del siglo v y primera parte del iv a. C., fecha, en general, reconocida para una buena parte de los opúsculos agrupados en el *Corpus*.

Los progresos de la etnología en la segunda parte del siglo xix y buena parte del xx han permitido iniciar una marcada tendencia a estudiar los fenómenos culturales de los llamados pueblos primitivos. La *Kulturhistorische Schule* (escuela histórico-cultural) ha fijado su especial atención sobre las etapas de la humanidad, y muy principalmente sobre las culturas primitivas superiores –egipcios, mesopotamios, indios, chinos– que se habrían constituido a partir de la fusión de tres tipos básicos: *a*) Un grupo formado por los pueblos cazadores y recolectores, cuyo rasgo cultural más típico es la magia y el empirismo. *b*) Otro grupo constituido por un estadio algo más avanzado, es decir, por culturas agrícolas, caracterizadas por el animismo. *c*) Pueblos ganaderos, organizados patriarcalmente, en cuya organización religiosa destaca el culto a un dios varón-padre.

¿Qué concepción presentan estos estadios culturales respecto al origen de la medicina, o mejor dicho, de la génesis de la enfermedad? En general, destaca la creencia básica de acuerdo con la cual el agente productor de la enfermedad es un dios o un espíritu, que envía al pecador, como castigo, la enfermedad. Dada la concepción religiosa de la enfermedad y de su origen, los medios de que dispone esta medicina pre-racional para curar la enfermedad es el *rito* (ensalmo, plegaria), si bien en no pocos casos rudimentarios tratamientos de orden dietético pueden acompañar a la mera plegaria.

Este tipo de medicina recibe el nombre de *credencial* y se caracteriza, básicamente, por una falta de adecuación entre el mal y el remedio. Dado el carácter empírico de tales culturas, pueden no faltar determinadas prescripciones aparentemente de carácter no religioso, y, sobre todo, en algunos casos asistimos a detalladas observaciones objetivas de los síntomas que acompañan a una dolencia. Por ejemplo, en algunos papiros egipcios de la época faraónica pueden leerse determinadas descripciones sintomatológicas, acompañadas de recetas y fórmulas mágicas. Entre estos papiros contamos con el llamado *Papiro Ebers* (redactado durante la dinastía XVIII y descubierto en 1862); el *Papiro Hearst* (de la misma época, descubierto en 1899); el *Papiro Chester-Beatty* (dinastía XIX). El *Papiro médico de Berlín* y el *Papiro médico de Londres* redondean la colección a la que estamos haciendo referencia.

Veamos, a guisa de ejemplo, un texto tomado del *Papiro de Berlín*. Se trata de un texto en el que se contiene la sintomatología de una enfermedad concreta, acompañada de prescripciones terapéuticas: "Pesadez de vientre; el cuello del corazón, enfermo; en el corazón, inflamación y latidos acelerados; las ropas le pesan al paciente y a veces no le proporcionan calor. Por la noche, sed. Mal sabor de boca, como si hubiera comido frutos de sicomoro. Carnes flojas, como las de los moribundos. Si desea evacuar, el vientre está hinchado y no se desahoga. De un enfermo así hay que decir que tiene una inflamación abdominal. Para librar al enfermo de la dolencia, hay que conseguir una contracción del abdomen: se aplicará un remedio para la inflamación, es decir, dátiles frescos, bayas de enebro, miel y frutos de ciruelo."

Frente a este documento, en el que no anotamos nada *supersticioso* en principio, he aquí un fragmento del *Papiro de Leiden* que contiene el conjuro para librar de una enfermedad: "¡Oh demonio que te alojas en el vientre de Fulano de Tal! Oh tú, cuyo padre se llama 'el que abate las cabezas', cuyo nombre es 'muerte', cuyo nombre es 'varón de la muerte', cuyo nombre es 'maldito para la eternidad'...!"

Está claro que esta medicina credencial, pre-racional, ofrece una serie de rasgos típicos: *1*) De un lado, la idea de que la enfermedad es una *mancha*, un pecado, y, por ende, que el remedio debe consistir en ritos catárticos o purificadores. *2*) Al lado de estos ritos catárticos, de origen mágico, destaca el hecho de que el médico egipcio posee no escasos conocimientos anatómicos e incluso fisiológicos.

La cultura *asirio-babilónica* ofrece rasgos parecidos a la egipcia en lo que atañe a la concepción nosológica. Una dolencia, una enfermedad, es entendida como una respuesta divina a un acto pecaminoso. El término empleado para designar ese concepto complejo culpa-enfermedad es *shêrtu*, que hace que el enfermo se sienta como un maldito de los dioses. El único tratamiento válido es el rito purificador.

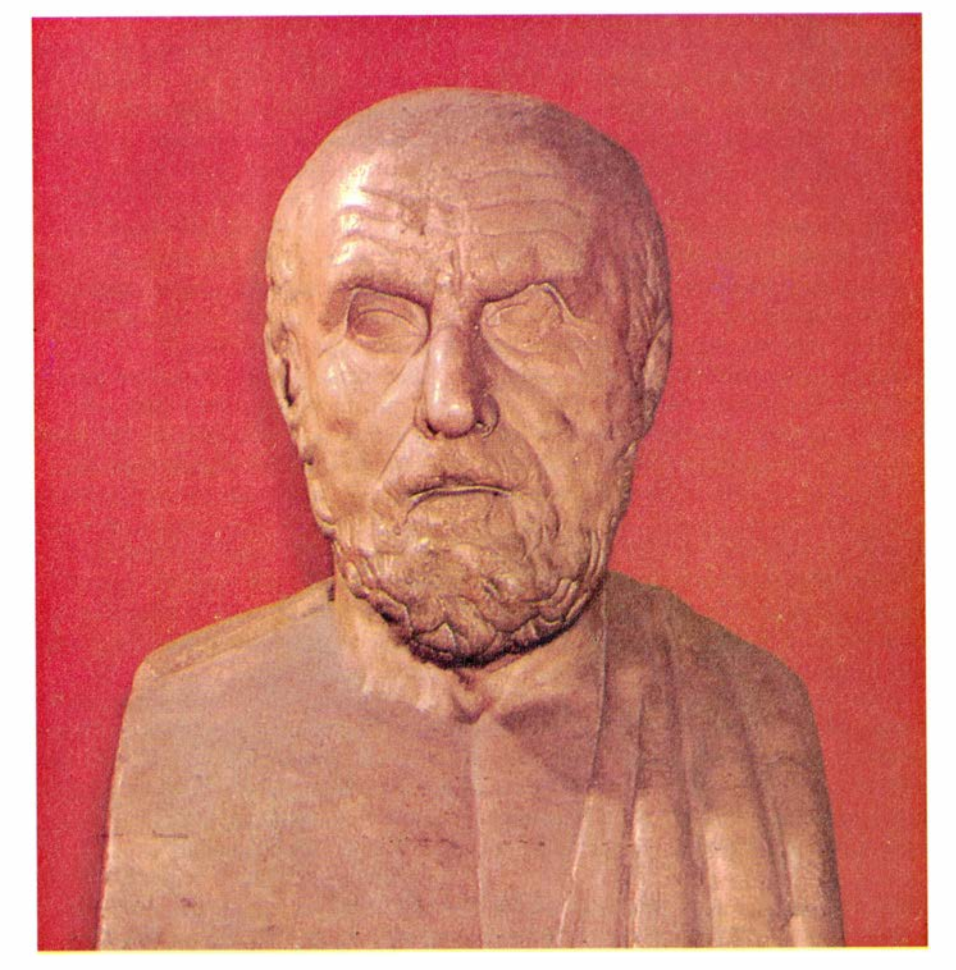
También Grecia ha conocido una

JURAMENTO HIPOCRATICO, cuya altura moral y deontológica admira, todavía, la medicina. Aunque estudios pormenorizados han demostrado que el texto, en su redacción definitiva, tal como lo tenemos hoy, es algo posterior a la época en que se supone que vivió Hipócrates (siglo v-iv a. C.), sin duda su espíritu reina en el texto. Lo fundamental del *Juramento* es la promesa de enseñar a sus discípulos el Arte (médico), considerar a su maestro y a sus hijos como miembros de la propia familia y el secreto. (La ilustración está tomada de un manuscrito bizantino del siglo xii, conservado en la Biblioteca Vaticana.)

etapa de medicina credencial. En los poemas homéricos, la noción de que los dioses pueden, y, de hecho, provocan una peste, se halla en la misma base de la *Ilíada*, donde Apolo envía una epidemia mortal a los griegos porque éstos han vejado a Crises, sacerdote del dios. Y también en este caso una serie de ritos catárticos consiguen calmar la cólera del dios (*Ilíada* I, 37-52).

Sin embargo, el caso más interesante de la persistencia, en el mundo griego, de una medicina credencial, la tenemos en el santuario de Asclepio, en Epidauro, donde las curaciones se realizaban, normalmente, mediante el rito de la *incubación*: el enfermo duerme en el recinto del templo; el dios le envía un sueño en el que se señalan las condiciones bajo las cuales la enfermedad puede curarse. Aquí, la ventaja es que disponemos de una serie de inscripciones votivas donde el paciente cuenta su propia experiencia. He aquí un ejemplo de estos *Yámata* o *Curaciones*: "Areta, lacedemonia, que era hidrópica: la madre de esta mujer dejó a su hija en Esparta y se dirigió a Epidauro, donde, mientras dormía en el templo, tuvo una visión: le pareció que el dios cortaba la cabeza de su hija y suspendía el cuerpo hacia abajo. Salió agua en abundancia; luego, el dios descolgó el cuerpo y volvió a unirlo a la cabeza. Después de esta visión, la madre regresó a Esparta, donde halló curada a su hija, y supo que ésta había tenido la misma visión." Otro curioso caso: "Agestrato y sus dolores de cabeza: este hombre sufría continuos insomnios debidos a sus dolores de cabeza. Una vez llegado al dormitorio (del templo), quedóse dormido y tuvo una visión: le pareció que el dios curaba su dolor de cabeza y le mostraba los ejercicios del pancracio. Al amanecer quedó curado y poco después venció en la prueba del pancracio en los Juegos Nemeos."

En verdad, estas famosas curaciones de Epidauro plantean sutiles problemas de interpretación, que no podemos abordar aquí. K. Kerényi, prestigioso historiador de la religión griega, ha dicho, de tales curaciones, que son milagrosas sólo en la medida en que toda curación, todo final feliz de una situación que implica la posibilidad de un final desastroso, es una especie de milagro. Pero hay importantes problemas psicológicos, y de otra índole, planteados por las *Curaciones* de Epidauro, que todavía no han recibido una respuesta adecuada.



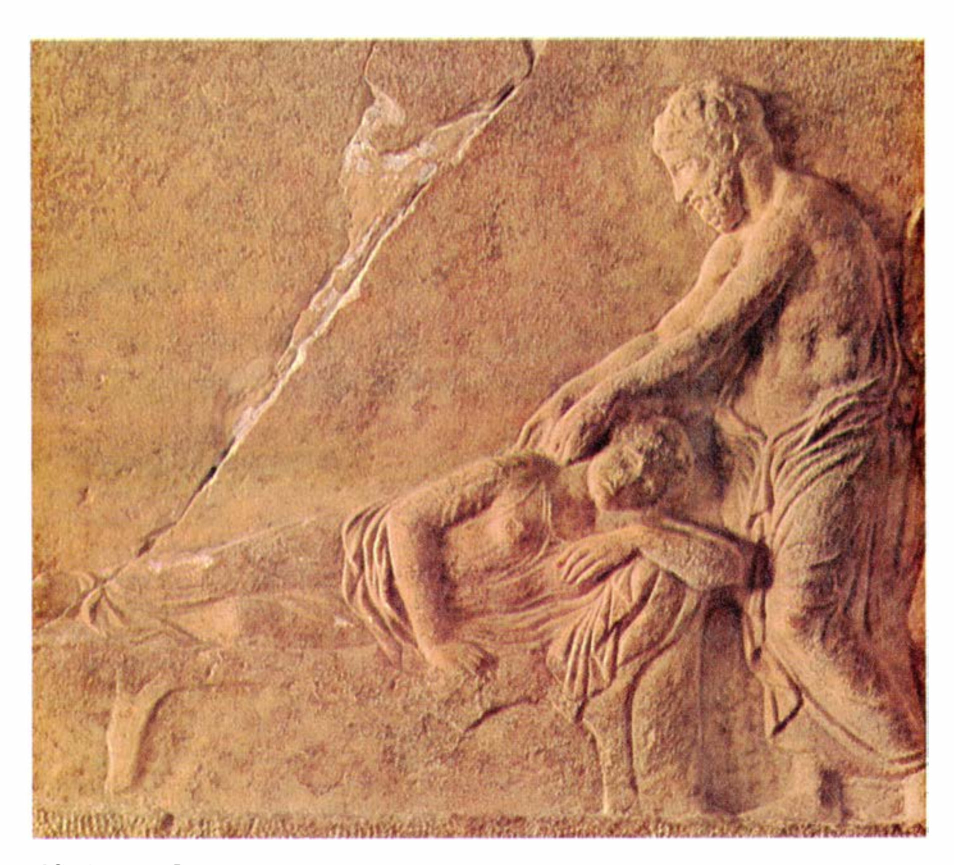
BUSTO DE HIPOCRATES, el famoso creador de la escuela médica de Cos y de una auténtica dinastía médica: su hijo y su nieto fueron destacados médicos, sobre todo su nieto; su yerno, Pólibo, es considerado autor del tratado *Sobre la naturaleza del hombre*. A principios del s. XX se dudó de su existencia real.

Como tendremos ocasión de comprobar a lo largo de este trabajo, en un momento decisivo de la historia griega se produce una fuerte reacción contra esa medicina credencial, representada, precisamente, por la medicina hipocrática, que reaccionará fuertemente contra la concepción divina de las enfermedades y sus medios de curación. Pero lo importante es señalar que, pese a todo, durante toda la historia de Grecia correrá una corriente irracional que la obra de Hipócrates y sus discípulos y colegas no logrará arrinconar. Tanto en la literatura como en la vida diaria, los *milagreros* representarán un elemento importante. De un lado, persistirá, sobre todo en la época arcaica, la concepción de la enfermedad como mancha, con una terapia especial a base de ritos catárticos o purificadores. En el santuario de Apolo en Delfos, pero también en Atenas, la práctica de la purificación será importante, y un buen ejemplo lo tenemos en la historia de Orestes, asesino de su madre y que es purificado por Apolo, tal como Esquilo lo ha plasmado en su tragedia *La Orestiada*. El exorcismo, el ensalmo (*epodê*) y la acción sanante de la música serán constantemente elementos curativos hasta bien avanzada la Antigüedad. Y, especialmente a lo largo del siglo VI a. C., anotaremos la presencia de famosos milagreros (Abaris, Epiménides, Aristeas), cuyas prácticas se ha querido relacionar recientemente con el chamanismo (D. R. Dodds). De ellos ha dicho Erwin Rohde: "No podemos llamarlos filósofos, ni siquiera precursores de la filosofía griega. Caminan precisamente en una dirección contraria, hasta el punto que el mayor esfuerzo de la filosofía, la exigencia más apremiante en el proceso de emancipación de los espíritus, fue desde un principio fraguarse otros principios y buscar otras direcciones."

Si una de las grandes aportaciones del espíritu helénico a la historia de Occidente fue la creación del pensamiento filosófico, tenemos ahora que afirmar que la hazaña de la elaboración del pensamiento médico no hubiese sido posible sin la especulación de los presocráticos. Filosofía y medicina serán, pues, dos de las más grandes contribuciones de Grecia a la creación de Occidente. Los jonios dieron un paso importante –pues una y otra aportación es fruto de la cultura jónica– en la historia espiritual de Occidente al iniciar

su lucha contra el mito y la concepción mítica del mundo. Si, de un lado, la especulación presocrática fue una superación de la etapa mítica de la concepción del mundo, la actividad médica significa la superación de la etapa pre-racional de la concepción de la etiología morbosa.

En efecto, en las costas de Jonia, en el litoral de Asia Menor, donde confluían culturas milenarias, tuvo lugar el paso de la cultura pre-racional arcaica a la cultura racional helénica. De este modo, el espíritu jónico aparece dotado de dos rasgos específicos y, al tiempo, complementarios: por una parte, una fuerte capacidad de abstracción que le permite llegar más allá de los meros fenómenos. Fue así como Jonia creó la metafísica, o algo que, andando el tiempo, llamaremos con ese nombre. Pero, simultáneamente, los jonios, pueblo práctico y emprendedor, supo habérselas con la realidad más inmediata y contar con ella. De ese segundo rasgo del espíritu jónico nacerá la etnología, y, con ella, la preocupación por lo que sucede en el cuerpo humano. Es decir, la medicina.



ESCENA DE *INCUBATIO*, procedimiento de curación milagrosa, practicada especialmente en el santuario de Asclepio en Epidauro. El enfermo, en esta práctica, se acostaba en el recinto del templo, y al dormirse, el dios, a través de un sueño, indicaba la terapia. La medicina científica no borró la credencial.

Pero esta medicina sólo será posible gracias a la especulación presocrática sobre la naturaleza, la *physis*. Más aún: el concepto mismo de *physis*, fundamental para entender la medicina helénica (todavía la palabra *fisiología* nos lo recuerda), no puede comprenderse del todo sin entrar en contacto con la *physiologia* presocrática. Lo ha dicho con palabras certeras un fino estudioso de la medicina y la filosofía antiguas como W. Jaeger: "La medicina –afirma– jamás habría llegado a convertirse en una ciencia sin las indagaciones de los primeros filósofos jónicos de la naturaleza, que buscaban una explicación natural de los fenómenos."

Al elaborar su doctrina de la *physis* el pensamiento griego sienta, pues, las bases de una profunda revolución en el campo de lo que será la medicina. Y entre las nociones presocráticas de la *physis* podemos señalar, con Pedro Laín, la universalidad, la necesidad y la racionalidad. "Cada cosa tiene su *physis* y esto indica que hay una *physis* común para todas las cosas." Es, diríamos nosotros, la *Physis* en mayúscula.

Si en un principio *Physis* fue una concepción "religiosa" de la naturaleza, esta noción se desacralizó en el pensamiento presocrático. Por un lado, Anaximandro sentó el principio de la regularidad de esta *physis*; en su concepción, los fenómenos de la *physis* ya no pueden atribuirse a un agente mítico, ni concebirse como una fuerza arbitraria. Aplicado al campo de la medicina, se sienta, por un lado, el principio de la *naturalidad* del hecho morboso, al tiempo que se traspone al cuerpo humano la misma noción de *physis*. En la medicina hipocrática se hablará, a partir de ahora, de la *anthrôpeia physis*, que, como la *physis* universal, está sometida a unas leyes. Con ello la enfermedad dejaba de ser un castigo divino para convertirse en un "desarreglo" de la naturaleza.

Junto a las aportaciones de Anaximandro, la de Empédocles, que introduce en el campo de la *physiologia* el principio de los cuatro elementos que constituyen la realidad. Y al aplicarse esta doctrina, surge la noción de los *humores*, que en el momento culminante de la medicina hipocrática son también cuatro: pituita, bilis amarilla, bilis negra y sangre. Tal doctrina la hallamos aplicada sistemáticamente, por vez primera, en el tratado hipocrático *Sobre la naturaleza del hombre*. Asimismo cabe señalar la profunda influencia de otros presocráticos en los orígenes de las doctrinas médicas: Anaxágoras, Demócrito, Diógenes de Apolonia, junto a una profunda impronta pitagórica.

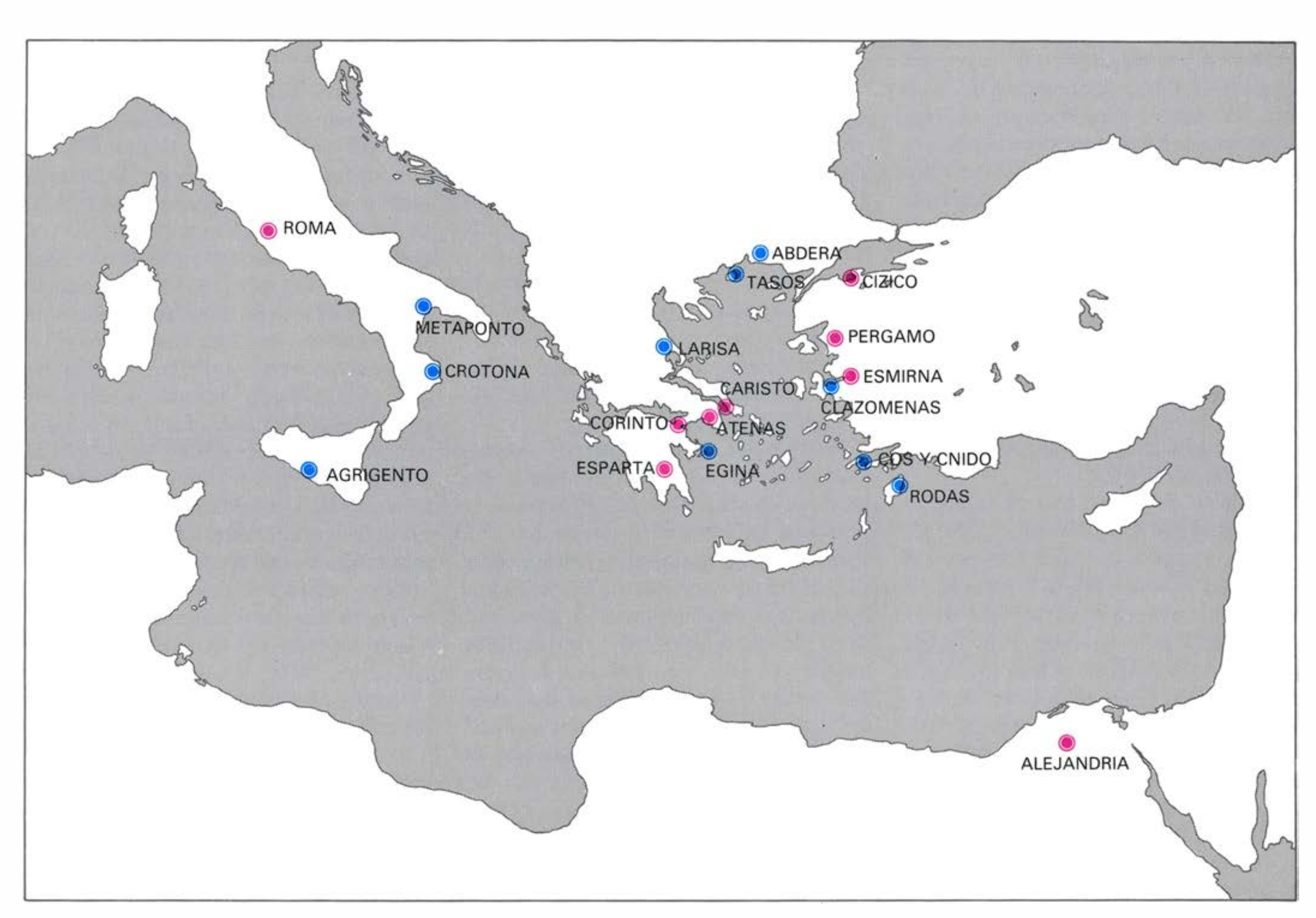
Sin embargo, cuando planteamos el problema del influjo de la filosofía sobre la medicina hipocrática, cabe hacer una clara distinción: de un lado, aquellos principios básicos que informan lo esencial de la ciencia médica. Así, según veíamos, la doctrina de la regularidad de la naturaleza (Anaximandro), la de los elementos (Empédocles) y otros que en este momento son menos importantes. De otro lado, la influencia particular de determinados pensadores presocráticos, que han aportado doctrinas concretas que no hallamos compartidas por todos los autores que forman el *Corpus* hipocrático. Cabe así señalar la presencia de elementos democriteos (estudiados por Wellman); el influjo concreto de Diógenes de Apolonia, sobre todo en la doctrina del *pneuma*; y, siguiendo esa misma orientación, se ha podido hablar del influjo de Heráclito en algún que otro opúsculo. También se han hallado influencias de la escuela pitagórica (Alcmeón de Crotona, posiblemente un pitagórico) e incluso de los sofistas.

Si hasta aquí hemos intentado ofrecer un cuadro de los antecedentes históricos y culturales de la medicina griega, ahora hemos de procurar un esbozo de las escuelas que, desde los tiempos de los que tenemos noticia, aparecen en el horizonte de la medicina griega. Porque la medicina helénica, en sus comienzos –y eso se extiende a toda su

larga historia– no ha conocido una sola orientación. Si en el siglo v podemos decir que las dos orientaciones más importantes son las contenidas en el *Corpus* hipocrático –la escuela de Cos y la de Cnido–, tenemos noticias de otras, como la de *Crotona*, fundada probablemente por Democedes –posiblemente el médico más antiguo conocido–, la de *Egina*, fundada, al parecer, también por Democedes; la de Sicilia, creada por Empédocles y cuyos representantes básicos fueron Acrón y Filistión; la de *Atenas*, en la que destacaron Mnesiteo y Dieuques, y la de *Rodas* y *Cirene*. Sabemos, asimismo, que hubo una orientación pitagórica. Alcmeón de Crotona parece que se relacionó con ella.

Pero, como decíamos, las dos escuelas más famosas del siglo v a. C. fueron las de Cos y de Cnido. Por lo menos, es de estas dos escuelas de las que nos han llegado más abundantes escritos. Precisamente el llamado *Corpus hippocraticum*, que reúne un material diverso y procedente, al parecer, de épocas muy distintas, se caracteriza por contener, al lado de escritos de carácter sofístico, otros de orientación claramente médica, que podemos dividir, *grosso modo*, en escritos de tendencia *empírica*, que suelen atribuirse a la escuela de Cnido (cuyo fundador fue Eurifonte, y en la que destacaron espíritus como Heródico de Selimbria, Ctesias, Alcamenes y otros) y otros opúsculos, de contenido y orientación más racional y científica, que suelen atribuirse a la escuela de Cos, cuya gran figura es Hipócrates.

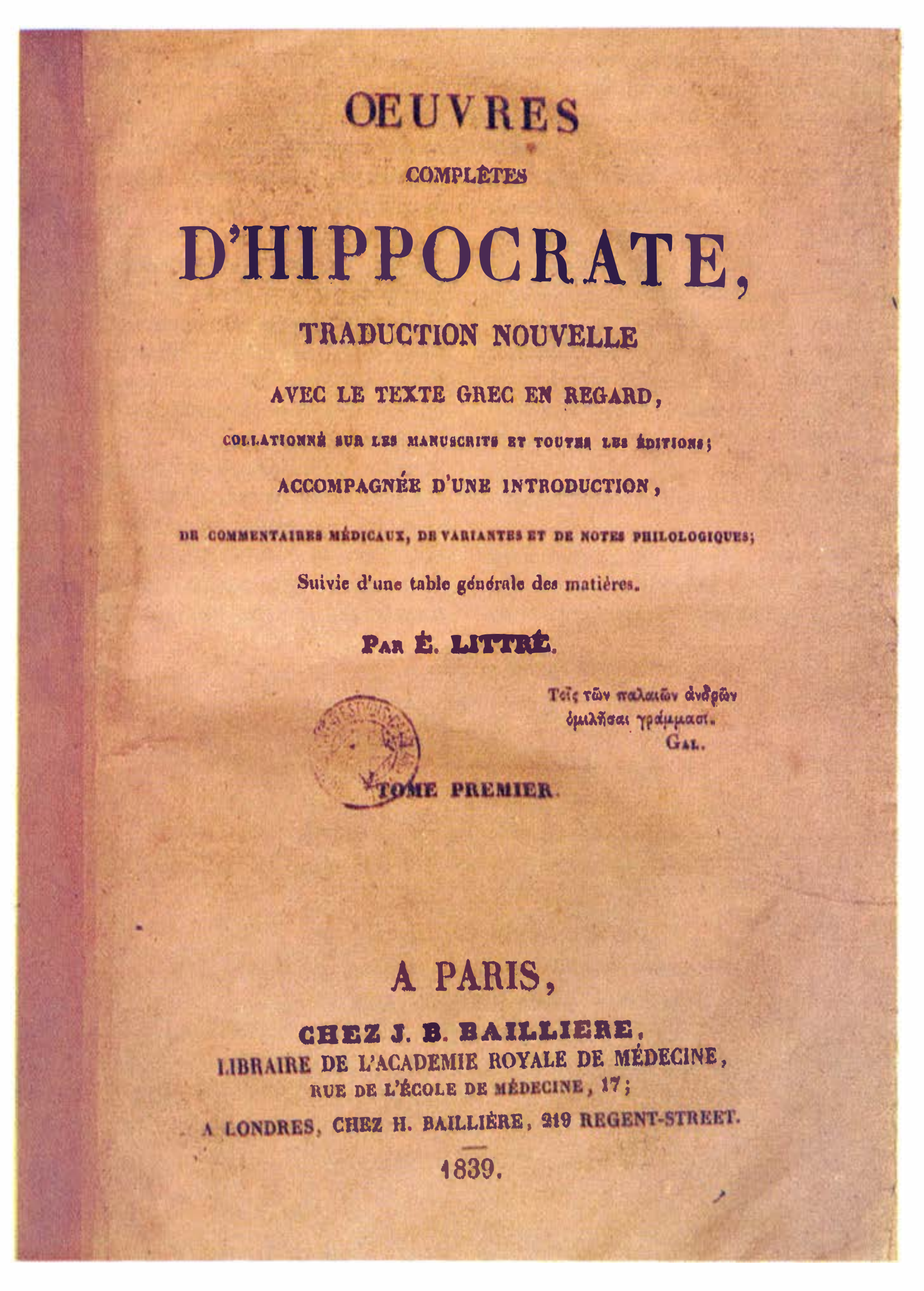
El gran problema que tiene planteado la historia de la medicina hipocrática es, empero, el de determinar si, de entre los escritos del *Corpus* atribuidos a la escuela de Cos, pueden señalarse algunos que, sin duda, hay que atribuir al gran maestro o si, como se afirmaba a comienzos del siglo xx, Hipócrates es un nombre detrás del cual no hay ningún escrito (Wilamowitz). Es la llamada *cuestión hipocrática*. No se trata, en esta discutida cuestión –como es el caso de la homérica– de si Hipócrates es una figura histórica real. Hoy no se duda en absoluto de la existencia real e histórica de Hipócrates, así como del espíritu que inyectó en su escuela. Sí, en cambio, se discute, aún hoy, si conservamos obras que le pueden ser atribuidas. Que el *Corpus* hipocrático es una mezcla confusa de escritos –quizá reunidos en la biblioteca alejandrina– lo reconocen los mejores especialistas. Así A. Rey ha afirmado taxativamente que “la colección hipocrática no es la obra de un solo hombre, ni de una sola escuela ni de una sola época” y Bourgey ha calificado el *Corpus* de “selva virgen”. Sin embargo, se han realizado importantes esfuerzos hermenéuticos por asignar al Maestro algunas obras, intentando con ello superar la posición hipercrítica del siglo pasado (Wilamowitz) que afirmaba no conocerse tratado alguno atribuible a Hipócrates (posición escéptica que comparte, hoy en día, el historiador de la medicina antigua L. Edelstein). Ya hacia finales de la segunda década del siglo presente O. Temkin señaló un posible camino, al proponer que la investigación debía orientarse hacia la búsqueda de lo que él llamaba la “conexión sistemática” de los diversos escritos del *Corpus*, con lo que al menos se puede señalar qué escritos pueden proceder de una misma mano. Siguiendo esta



PRINCIPALES CIUDADES de la antigüedad greco-romana que destacaron por la práctica de la medicina. No todas las escuelas de esas ciudades florecieron al mismo tiempo. Una de las más antiguas fue la de Crotona, patria de Alcmeón, cuyas doctrinas influyeron en las escuelas posteriores. En el siglo v destacan las de Cos y Cnido. Casi nada sabemos de las de Rodas, Egina y Atenas. Alejandría fue la ciudad más importante durante la época helenística.

postura se ha trabajado, sobre todo en Alemania, por parte de críticos como H. Diller y K. Deichgräber. En pleno siglo XIX, Littré ya había hecho notables esfuerzos, atribuyendo a Hipócrates los siguientes opúsculos del Corpus: *Antigua medicina, Pronóstico, Aforismos, Epidemias* I y III, *Régimen de las enfermedades agudas, Aires, aguas y lugares, Articulaciones, Fracturas, Instrumentos de reducción, Heridas en la cabeza, Juramento* y *Ley*. Otros críticos, menos decididos, distinguen un hipocratismo lato, estricto y estrictísimo (Laín).

La considerable diversidad de los escritos que constituyen el *Corpus* hipocrático se refleja, por lo pronto, en el aspecto, digamos, literario. De los sesenta y pico tratados que contiene, unos son escritos de relativa extensión. *Sobre el régimen* está formado por cuatro libros; los tratados traumatológicos (*Heridas en la cabeza, Articulaciones, Fracturas, Palanca*) son asimismo bastante extensos. Y lo son, también relativamente, una serie de tratados –posiblemente los más famosos– como *Antigua medicina, Pronóstico, Naturaleza del hombre, Aires, aguas y lugares*, y, en especial, los tratados ginecológicos. Al lado de estos opúsculos que abarcan un número importante de páginas, otros tratados se caracterizan por su brevedad: *Juramento, Ley, Humores, Dentición, Uso de líquidos*, etc. Pero en lo que realmente se distinguen es en su estructura formal: hay tratados que son casi obras de arte literarias, en las que el estilo y el lenguaje están muy cuidados. A su lado, tenemos ciertos opúsculos que son meras notas, sin gramática, sin sintaxis, como si de apuntes tomados en una conferencia se tratara: así, *Sobre el corazón y Humores*, por citar los más significativos. En otro orden de cosas, mientras ciertos tratados son meras exposiciones descriptivas, otros tienen un contenido polémico que los hace altamente interesantes, y que nos introducen, en cierto modo, en los enfrentamientos teóricos que, sin duda, tendrían lugar en la geografía griega entre los representantes de las distintas escuelas. En cuanto a las orientaciones teóricas de los distintos tratados, diremos que se han podido distinguir, a grandes rasgos, tres actitudes: la de un grupo de escritos que se caracterizan por su actitud empírica, y que hoy tienden a atribuirse a la escuela de Cnido. Otros delatan un espíritu en el que se combina la observación junto al razonamiento, y que pueden calificarse, en cierto modo, de positivistas. Son los escritos que suelen atribuirse a la escuela de Cos, si bien últimamente se va haciendo clara la idea de múltiples intercambios entre las dos tendencias. Finalmente, una serie de tratados que delatan a un autor no especialista, no médico, pero que toca, a modo de conferencia, un tema médico: así, entre otros, *Sobre el Arte y Ventosidades*, por ceñirnos a los más notables.


OEUVRES

COMPLÈTES

D'HIPPOCRATE,

TRADUCTION NOUVELLE

AVEC LE TEXTE GREC EN REGARD,

COLLATIONNÉ SUR LES MANUSCRITS ET TOUTES LES ÉDITIONS;

ACCOMPAGNÉE D'UNE INTRODUCTION,

DE COMMENTAIRES MÉDICAUX, DE VARIANTES ET DE NOTES PHILOLOGIQUES;

Suivie d'une table générale des matières.

PAR É. LITTRÉ.

Τοῖς τῶν παλαιῶν ἀνδρῶν
ὁμιλῆσαι γράμμασι.
GAL.

TOME PREMIER.

A PARIS,

CHEZ J. B. BAILLIÈRE,
LIBRAIRE DE L'ACADEMIE ROYALE DE MÉDECINE,
RUE DE L'ÉCOLE DE MÉDECINE, 17;
A LONDRES, CHEZ H. BAILLIÈRE, 219 REGENT-STREET.

1839.

PORTADA DE LA FAMOSA EDICION del *Corpus* hipocrático cuidada por E. Littré (1801-1881), famoso médico y filósofo francés, cuya labor en pro del conocimiento de Hipócrates fue muy destacada.

Pero el problema básico es éste: ¿cuál es el contenido científico de los tratados del *Corpus* hipocrático? Por lo pronto, cabe afirmar que en él se hallan contenidas las distintas especialidades médicas. Tenemos escritos de carácter general, patológicos, quirúrgicos, terapéuticos, traumatológicos, ginecológicos, obstétricos, pediátricos, clínicos, deontológicos, sin que ello quiera decir que, en la época hipocrática, la ciencia médica hubiese llegado a un alto grado de especialización. Ello no es así, y en algunas ocasiones los propios autores del *Corpus* lo señalan. Pero hay, en cambio, una doctrina general –con variantes, según los autores y las tendencias– y una conciencia muy clara de la misión de la medicina y de la actividad del médico.

Por lo pronto, se tiene clara conciencia de que, combinando tradición, observación, experiencia y razón (el *lógos* griego) se ha constituido un arte médi-

co, una *tékhnê*, que dispone de los requisitos intelectuales imprescindibles. Y esa *tékhnê* tiene una misión muy clara: la curación de la enfermedad. Por otro lado, esa *tékhnê*, como tal, puede enseñarse a los otros. El médico hipocrático no cree en el carácter sobrenatural de la enfermedad: toda dolencia tiene unas *causas*. Lo que varía, en los escritos, es la determinación de las mismas. Unas veces es el *desarreglo humoral*, otras el influjo del medio ambiente, otras, en fin, determinados agentes, como la famosa doctrina de las *perissómata* o residuos. Pero siempre el médico sabe que hay que detectar un agente concreto que ha producido el desarreglo patológico.

La doctrina básica de la *fisiología* hipocrática es la famosa tetralogía humoral: la naturaleza humana está constituida por cuatro elementos fundamentales, pituita, sangre, bilis amarilla y bilis negra (hay variaciones en cuanto al número y en cuanto a la esencia, pero debe tratarse de una evolución en el interior de las escuelas). Esta doctrina es expuesta por primera vez de un modo específico en el escrito *Sobre la naturaleza del hombre*, atribuida generalmente a Pólibo, yerno de Hipócrates. Del equilibrio de estos humores depende la salud, así como, sobre todo en la medicina posterior, la existencia de determinados tipos humanos según el predominio *temperamental* (sanguíneos, flemáticos, biliosos y melancólicos).

Importante en este orden de ideas es el papel que suele concederse al cerebro. Para la escuela de Cos, es el cerebro el centro básico del funcionamiento del cuerpo humano. En otras escuelas se asignará este papel al corazón (así la escuela siciliana, y, siguiendo a ésta, Platón).

Por las arterias circula *pneuma*, por las venas sangre. Esta doctrina, explicable por el hecho de que el médico hipocrático sólo podía trabajar sobre cadáveres, tardará mucho en ser desarraigada por la medicina griega. Un punto importante en la práctica médica hipocrática es la doctrina del pronóstico y el diagnóstico. El médico, mediante un adecuado interrogatorio del paciente y el análisis de los síntomas está en condiciones de determinar la causa de la dolencia y su solución futura. Hay que señalar, con todo, la diferencia profunda que existe entre el pronóstico hipocrático y el de la medicina moderna. Algunos críticos han pretendido negar todo valor científico al pronóstico hipocrático (Edelstein), pero otros son menos severos a la hora de establecer su carácter científico. Hay que añadir que la recopilación de datos lleva a una prehistoria de lo que será la historia clínica. En el tratado *Aforismos*, por ejemplo, estamos ante un escrito en el que no hay reflexión, sino mera recopilación de datos para uso del futuro médico. Lo mismo cabe decir de los libros sobre *Epidemias*.

Lo que más ha impresionado a la posteridad de la medicina hipocrática es su altísimo nivel deontológico. El famoso *Juramento* será siempre un modelo como código moral de la medicina.

Se ha afirmado que la dietética es el elemento más importante de la terapéutica hipocrática. Pero hallamos asimismo, aunque en un grado no muy desarrollado, una farmacología que alcanzará su pleno apogeo en etapas posteriores de la medicina griega, con Dioscórides.

Una vez esbozado el contenido general de la medicina hipocrática, creemos de cierto interés ocuparnos, aunque sea brevemente, de la constitución y elaboración de los estudios consagrados a Hipócrates y su actividad médica. Tras el renacimiento que significó la labor pionera de un Littré y un Petrequin, en la mitad del siglo XIX, se produce un cierto estancamiento de estos estudios. En general, cabe decir que este estancamiento corre parejo con la crisis general que acompaña a los estudios de historia de la medicina. En el campo concreto de la historia de la medicina griega, los últimos años del siglo XIX y los comienzos del XX van acompañados de una cierta actividad filológica e histórica, a la que van unidos los

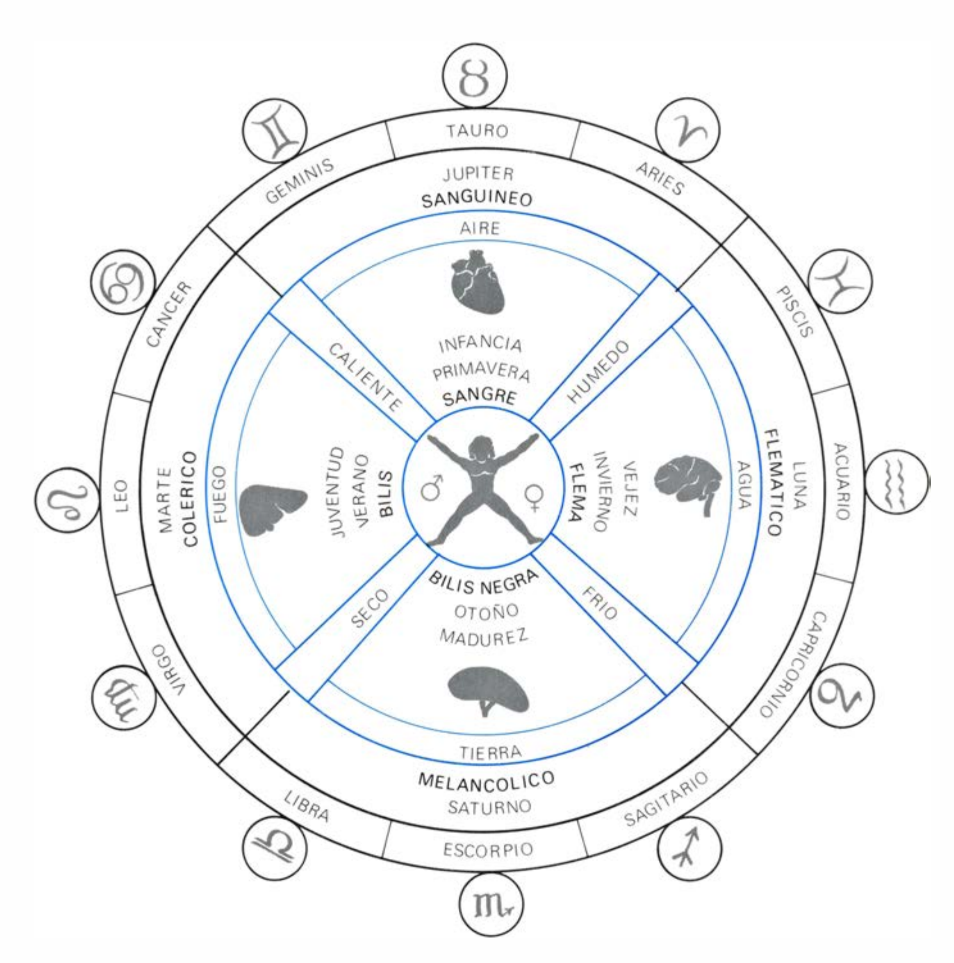


PROCEDIMIENTO DE LA REDUCCION DE LAS LUXACIONES, según se recoge en un manuscrito medieval hipocrático, que se conserva en la Biblioteca Nacional de Bolonia. La traumatología en Hipócrates es arcaica, pero en el Corpus hay varios textos que se ocupan de ella **[*véase tríptico de la página 27*].**

nombres de un Wellmann, de un Wilamowitz, de un Fredrich, un Diels y un Heiberg. Esta labor culmina, en 1901, con la elaboración de un plan de publicación del *Corpus medicorum graecorum*. Sin embargo, fue hacia el decenio de los años veinte cuando comprobamos nuevos impulsos, sobre todo en Alemania, hacia la constitución definitiva de una concepción histórica de la medicina, y, con ella, de la hipocrática y griega en general. El *Sudhoff Archiv für die Geschichte der Medizin*, junto con la revista *Kyklos* son los exponentes máximos de ese renacimiento. Los nombres fundamentales son ahora O. Temkin, W. Jaeger, L. Edelstein y Fr. Solmseb, quienes, sin embargo, sentirán pronto la grave amenaza que el control ideológico del nazismo esgrime contra sus actividades, que se patentiza en un texto del tercer número de la revista *Kyklos*: "Las perspectivas para el futuro son muy turbias."

La emigración de buena parte de ese grupo de investigadores a Norteamérica convierte, por un momento, a esta nación en el centro de gravedad de los estudios de historia de la medicina antigua. Simultáneamente, sobre todo en Francia, comienza a manifestarse un notable interés por los problemas de la medicina helénica, al tiempo que en Alemania aparece otra generación de estudiosos (H. Diller, K. Deichgräber) que pronto van a tomar el relevo de sus emigrados maestros. Asistimos así, hacia los años treinta, a un momento de auge en este campo de la investigación, cuyo exponente es la continuación del *Corpus medicorum graecorum*, al lado de importantes revistas consagradas al tema: el *Bulletin of the History of Medicine* y la continuación del *Sudhoff Archiv*.

En el momento actual, cabe señalar una auténtica floración de estudiosos que aúnan sus esfuerzos para ampliar y profundizar nuestros conocimientos. En Alemania, figuras como Grensemann, Michler, Flashar continúan la línea de sus maestros; en Francia, Jouanna, Robert; en Suiza, W. Müri, Kudlien; en Italia Perazzi, Vegetti, G. Lanata; en Inglaterra, Lonie y Lloyd; en España, P. Laín, a la cabeza de un grupo de estudiosos, representa la aportación castellana, y, a su lado, el esfuerzo de la escuela catalana (E. Vintró y el autor de estas líneas), atestiguan un momento importante en la historia de la investigación; en Bélgica R. Joly ha iniciado la publicación de una nueva edición del *Corpus* hipocrático dentro de la colección G. Budé.



CORRESPONDENCIA ENTRE LOS CUATRO HUMORES (pituita, bilis amarilla, bilis negra y sangre), las cuatro estaciones del año, las cualidades y los cuatro elementos (fuego, aire, agua y tierra).

Al lado de esta actividad investigadora, la aparición de una nueva corriente valorativa del espíritu de la medicina hipocrática, que ha cristalizado en el movimiento neohipocrático, es un fenómeno digno de ser considerado. No se desconoce el profundo abismo que separa la moderna medicina de los principios técnicos, más bien pobres, que caracterizan al hipocratismo. Pero se reconoce que, en las páginas del *Corpus* hipocrático, late una grandeza ética y deontológica que merece ser conservada.

Pese a balbuceos iniciales de esa orientación (representados en la obra de W. Kraus, *Patología general y especial de la persona*, aparecida en 1919), la verdadera historia del neohipocratismo se hace realidad plena en 1938, fecha del primer congreso nacional, celebrado en Marsella. Le siguen los cuatro primeros congresos internacionales de medicina neohipocrática (París, 1937; Evian, 1953; Montecatini, 1956; Atenas-Cos, 1958; y Madrid, 1965). Son perceptibles claramente, en este movimiento, dos líneas de orientación: de un lado, el representado por los historiadores (Castiglioni, Pazzini, Bourgey, Laín), al que hay que añadir la larga lista de los filólogos que hacen, paralelamente, historia del hipocratismo; de otro, los clínicos, entre los que destacan los nombres de Cawadias, Martiny, Giraud, Guy-Laroche, Laignel-Lavastine y Aschner.

A modo de justo complemento de estos congresos, conviene señalar la existencia de los simposios de medicina hipocrática, acaparados fundamentalmente por filólogos, aunque participen en ellos notables historiadores de la medicina griega. Iniciados en Estrasburgo en 1970, se han celebrado ya otros en Mons (1974) y París (1978). El próximo tendrá lugar, este año, en Lausanne. Se congregan aquí los mejores investigadores de la medicina hipocrática, y se discuten los problemas básicos que plantea su estudio. Es digno de tenerse en cuenta que las actas son pulcramente editadas, lo que permite al interesado acceder a la problemática básica y a las orientaciones que determinan este campo de investigación.

Nos queda, finalmente, un aspecto importante por tratar: ¿cuál es el *nivel intelectual* de la medicina griega

hipocrática? La pregunta no nos la formulamos gratuitamente. En 1938, G. Bachelard publicaba un interesante libro (*La formation de l'esprit scientifique*) en el que, partiendo de lo que él llama los obstáculos epistemológicos, analizaba la paulatina elaboración, a través de tres etapas, del espíritu científico. Para G. Bachelard, y precisamente de acuerdo con el distinto nivel intelectual conseguido a base de un "psicoanálisis del conocimiento objetivo", podemos distinguir tres grandes momentos en esta importante evolución del conocimiento: un período precientífico, que llega hasta el siglo XVIII; un período científico, que comprendería todo el siglo XIX; y, finalmente, el nuevo espíritu científico que se inicia a principios del siglo XX con la teoría de la relatividad.

Aunque Bachelard se ocupa, en su libro, fundamentalmente del siglo XVIII, resulta implícito en su trabajo que toda la producción anterior a este siglo cae bajo la misma denominación de *precientífica*. Con lo que la ciencia griega, y con ella la medicina, queda englobada bajo la misma rúbrica. Sin embargo, algunos trabajos de autores más o menos inspirados en las doctrinas del tratadista francés han intentado, con mayor o menor fortuna, abordar la producción helénica para confirmar sus puntos de vista. El más importante –y quizás el único llevado a término con relativa extensión– es un estudio del profesor belga R. Joly en un libro que merece la pena analizarse (*Le niveau de la science hippocratique;* París, 1966).

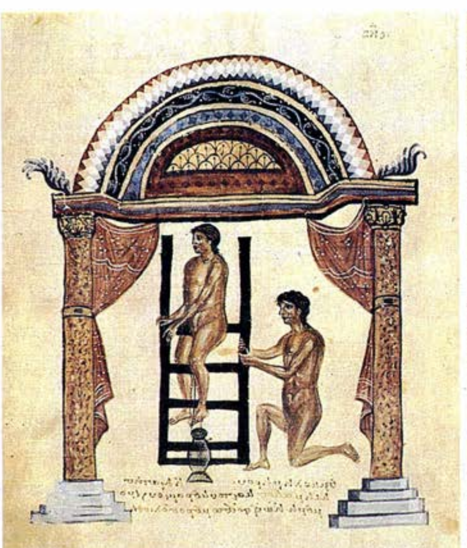
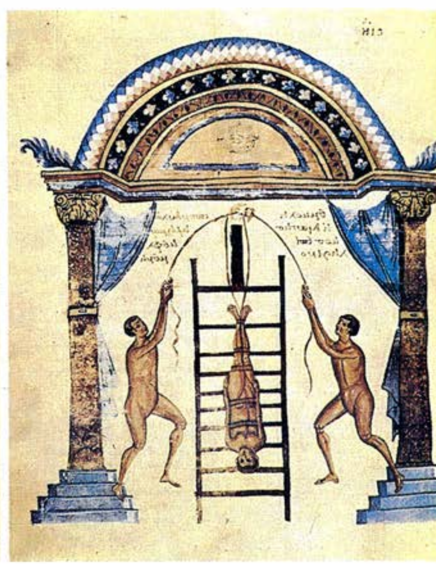
El punto de partida de la obra de R. Joly es el hecho de que, en la valoración de la figura de Hipócrates, asistimos a una doble postura: de una parte, la de aquellos críticos, a quienes podríamos llamar "modernistas", que ofrecen un saldo positivo de las aportaciones médicas de la medicina de la escuela de Cos y de Cnido recogidas en el *Corpus* hipocrático, y que consideran que la actitud científica de la medicina griega antigua puede perfectamente ponerse en paralelo con la medicina moderna. Sólo una diferencia de *cantidad*, no de *calidad* las distinguiría. En esta misma línea, con puntos de vista más o menos diversos, estarían figuras como G. Senn, Bourgey, Perazzi, Martiny, Kanatsoulis, Lichtenhaehler, ya en el siglo XIX representada esta orientación en autores como Littré y Pétréquin. Para estos autores, el médico hipocrático trabaja con orientaciones perfectamente "modernas": realiza observaciones de la naturaleza y de la enfermedad, practica la experimentación, y, en suma, adopta una auténtica actitud científica ante el hecho morboso. Contra esta valoración –que en el caso de Kanatsoulis adquiere sesgos casi caricaturescos– está el punto de partida de Bachelard, que quiere situar la ciencia médica griega en su verdadero estadio, en su nivel intelectual.

Joly lleva a término, en su trabajo, un análisis minucioso de los diversos obstáculos epistemológicos bachelardianos que pretende descubrir la medicina hipocrática: polivalencia causal (una misma causa puede producir varios efectos), sobredeterminación (una causa puede verse complementada con otra), polifarmacia (para el tratamiento





EL *CORPUS* HIPOCRATICO, como las obras de los demás médicos griegos, fueron traducidas por los árabes, y, más tarde, en Occidente, al latín. El cuadro de la izquierda muestra una página de una versión latina de los *Aforismos*, y el de la derecha, otra página de una versión árabe de Galeno.

DIFERENTES PROCEDIMIENTOS HIPOCRATICOS para reducir luxaciones. Los principales tratados traumatológicos del *Corpus* hipocrático son: *Heridas en la cabeza, Fracturas, Articulaciones, Instrumentos de reducción*. El rasgo que caracteriza la reducción traumatológica es que ésta se hace de acuerdo con la posición natural del órgano afectado. El *Corpus* contiene sesenta y pico tratados, de extensión, contenido y calidad literaria muy dispares.

de una sola enfermedad se aconsejan diversos tipos terapéuticos), el llamado obstáculo substancialista (valor especial de determinadas *substancias*, por ejemplo, piedras preciosas), animismo (concepción de los agentes morbosos como auténticos seres que actúan como dotados de vida), libido (papel excepcional de los órganos sexuales), apriorismos (misoginia, papel de la derecha frente a la izquierda), etcétera. Una vez ha llevado a término este minucioso análisis de los principios básicos que informan la medicina hipocrática, Joly proclama que esta medicina cae, perfectamente, bajo la denominación de *precientífica*.

Sin negar, de entrada, el valor del análisis de Joly, cabe empero establecer determinadas consideraciones: *1*. Es cierto, a grandes rasgos, que existen, en la medicina, e incluso en la ciencia griegas, determinados obstáculos epistemológicos: hallamos en ella claros indicios de apriorismo. Por ejemplo, y limitándonos al campo de la pura medicina, la idea de que el feto femenino tarda algo más que el masculino en formarse; que hallamos en esta medicina creencias animistas que pueden detectarse en la concepción de la enfermedad como un ataque de ciertos agentes morbosos; que hay un notable deductismo, importantes concepciones teleológicas y claros indicios de que la observación e incluso la experimentación que realiza el médico hipocrático están viciados por puntos de partida establecidos *a priori*, lo que falsea las mismas observaciones. Esto es claro, por ejemplo, en las ideas genéticas que hallamos en el tratado *Sobre la generación y la naturaleza del niño*.

2. El fuerte influjo de la filosofía sobre las ideas del médico hipocrático, detectables en no pocos tratados del *Corpus* (papel del aire en la etiología morbosa, obvia influencia de Diógenes de Apolonia; claros indicios de influjo de Demócrito, de Empédocles y de Anaxágoras, han sido señalados también en el *Corpus*), no llega a compensar la tendencia a presentar la medicina como un campo propio e independiente de la filosofía tal como podemos hallarla en la dura polémica que realiza el autor del tratado *Sobre la medicina antigua*. Eso redundó en una mayor tendencia a la deducción frente a la inducción y al empirismo, propios de la medicina moderna.

3. Pero aun aceptando los dos puntos anteriores, es preciso señalar que la medicina hipocrática presenta unos rasgos importantes que la separan claramente de la medicina de los pueblos del próximo Oriente: por lo pronto, y frente al carácter credencial de la medicina pregriega, no puede olvidarse que la medicina hipocrática (véase el tratado *Sobre la enfermedad sagrada*) significa el fin de la concepción de la enfermedad como algo causado por un dios. Con Hipócrates la medicina se convierte en un estudio de la naturaleza humana, con sus leyes y sus principios. Junto a este trascendental hecho, la medicina hipocrática estableció, de una vez por todas, la regularidad de esta misma naturaleza humana, su esencial constitución fisiológica (la teoría de los *humores*, claramente expuesto en el tratado *Sobre la naturaleza del hombre*); el importante valor que, entre otros, se concede a la combinación de experiencia y reflexión en el tratado *Preceptos*; la decisiva aportación a la medicina universal del concepto de pronóstico (y eventualmente del diagnóstico) aunque en algunos casos estos conceptos no equivalgan del todo a sus correspondientes modernos; el decisivo descubrimiento que significa la relación medio ambiente-enfermedad, clarísimamente establecido en el tratado *Sobre los aires, aguas y lugares*; finalmente, la noción de historia clínica tal como aparece practicada en los distintos libros sobre *Epidemias*.

¿Cuál es la conclusión a que podemos llegar tras estas consideraciones? A nuestro juicio, a la siguiente: si aceptamos con Joly-Bachelard que puede dividirse la formación del espíritu objetivo en tres etapas (precientífica, científica y nuevo espíritu científico) cabe, por otra parte, trazar claras subdivisiones en el período precientífico; que mientras la medicina prehelénica caería dentro de la denominación de medicina credencial, la medicina hipocrática debe ser llamada, sin ambages, una medicina racional, en la que muchos aspectos precientíficos y creencias puedan subsistir como fósiles más o menos vivos, pero en la que apunta una nueva orientación: la de la experiencia contrastada con la razón (*tribê metà lógou*). Que, en suma, el nivel intelectual de la medicina hipocrática es un nivel lógico (*logos*), racional; que no llega a la geometrización propia de la ciencia barroca (Kepler, Galileo), pero que, en parte, la preanuncia.

Reparación inducible del ADN

La lesión del material hereditario desencadena la "respuesta SOS", en la que varios sistemas enzimáticos actúan coordinadamente. Dos proteínas, RecA y LexA, desempeñan múltiples funciones en la reparación del daño

Paul Howard-Flanders

La vida de un organismo y su continuidad de generación en generación dependen de la estabilidad a largo plazo de la información genética cifrada en la doble hélice de ADN. Estabilidad que no debe traducirse por invulnerabilidad a las lesiones. El ADN, sea de una célula bacteriana o de los millones de células que componen un ser humano, se muestra muy sensible a los daños producidos por las radiaciones perjudiciales y por los agentes químicos presentes en el medio ambiente. Su estabilidad y la precisión de su replicación están protegidas por enzimas que continuamente reparan las lesiones genéticas. En concreto, las lesiones importantes de la molécula de ADN pueden inducir una respuesta de emergencia mediante la cual se sintetizan mayores cantidades de enzimas reparadores, que actúan en perfecta coordinación. Durante mucho tiempo se vino observando y estudiando varios sistemas de reparación del ADN, pero ha sido en los dos últimos años cuando han empezado a comprenderse los pormenores de la naturaleza de algunas de las respuestas inducidas y los sistemas reguladores que las controlan.

Tanto la vulnerabilidad a las lesiones como la susceptibilidad a la reparación van implícitas en la misma estructura del ADN. Las cadenas de la doble hélice son una sucesión de moléculas de azúcar y grupos fosfato. De este esqueleto sobresalen una serie de grupos químicos llamados bases. La unidad constitutiva real del ADN son los nucleótidos, cada uno de los cuales consta de una base, un azúcar y un grupo fosfato. Las bases son de cuatro tipos: dos purinas, adenina (*A*) y guanina (*G*), y dos pirimidinas, citosina (*C*) y timina (*T*). Las bases de una cadena se unen por enlaces de hidrógeno a las bases de la otra cadena, para formar el ADN bicatenario. Además, las bases son complementarias: *A* se aparea sólo con *T*, y *G* sólo con *C*. Se sigue de ello que las dos cadenas son complementarias y que, efectivamente, cada una de ellas es un molde de la otra: si una cadena tiene la secuencia *TGACCGCTTA*, la otra cadena deberá leerse *ACTGGCGAAT*. La complementariedad constituye el fundamento de la replicación del ADN, de su expresión en forma de proteína y de su recombinación. Durante la replicación, cada cadena sirve de molde para fabricar una cadena "hija". En la expresión, la información hereditaria cifrada en la secuencia de bases, de una de las cadenas, que componen un gen se transcribe en una secuencia complementaria de una cadena de ARN, que posteriormente se traducirá a proteína; cada tres bases, un codón, determinan uno de los aminoácidos que integran la cadena de proteína. En la recombinación genética, las cadenas de ADN se rompen y las regiones homólogas se unen para dar lugar a nuevas combinaciones de genes.

Según esto, la incorporación en el ADN de una base incorrecta o alterada, así como la presencia de una lesión que distorsione la doble hélice o impida el perfecto apareamiento de las bases, obstaculizarán la replicación, la síntesis de proteínas y la recombinación. A menos que se repare satisfactoriamente la lesión causada al ADN puede sobrevenir la muerte de la célula; parece que en las células de mamífero tales lesiones pueden a veces iniciar un proceso que conduce al cáncer.

De muchas maneras se puede atentar contra la integridad del ADN. Sin embargo, la lesión cuya reparación ha sido objeto de mayor atención es el dímero de pirimidina, producido por la exposición a radiaciones ultravioleta. Se forma un dímero cuando dos bases pirimidínicas (timinas o citosinas) adyacentes en una cadena se unen para constituir un anillo de cuatro carbonos. Las dos bases que forman el dímero pierden la alineación correcta, rompiéndose los enlaces de hidrógeno que las unen a las bases complementarias y distorsionando el esqueleto del ADN; esto impide, a su vez, el apareamiento debido de dos bases a cada lado del dímero. Basta la presencia de un solo dímero de pirimidina para interrumpir la transcripción y la replicación.

Los conocimientos fundamentales sobre el proceso de reparación del ADN proceden, en gran parte, de experimentos genéticos realizados con células bacterianas, expuestas a radiaciones ultravioleta, y de estudios bioquímicos sobre la replicación y reparación del ADN y de los enzimas que participan en esos procesos. En este artículo se repasará cronológicamente la historia de algunos de los descubrimientos, advirtiendo al lector (y esperando que lo comprenda) que los sistemas biológicos complejos no suelen resolverse en su totalidad de una sola vez, ni siquiera en una progresión lógica. Por el contrario, es normal que se acumulen observaciones y datos sin una relación obvia o una significación inmediata; los datos van adquiriendo valor y, gradualmente, las piezas encajan y empiezan a cobrar sentido.

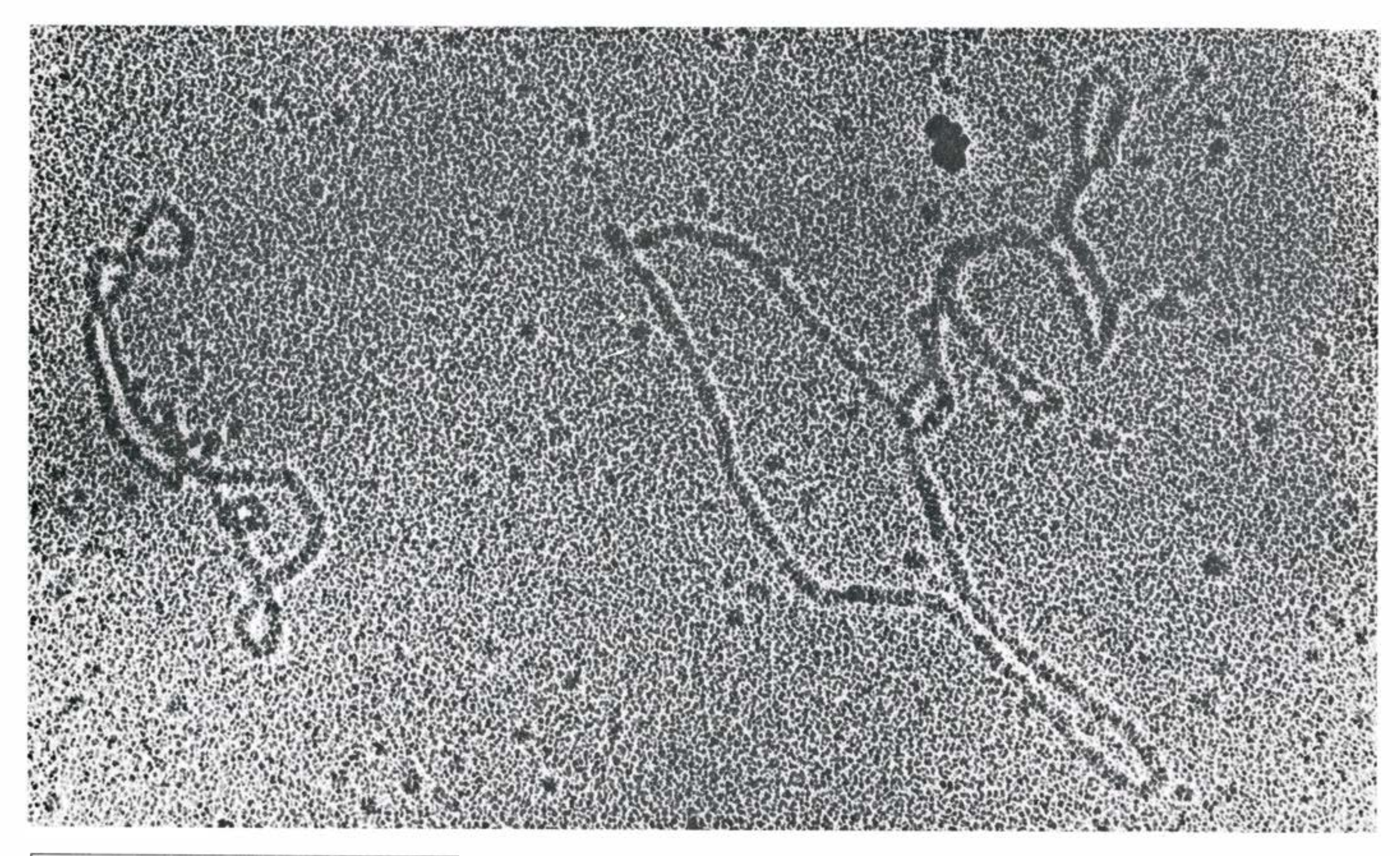
A finales de la década de 1950 se sabía que la bacteria intestinal *Escherichia coli* sobrevivía, en general, a exposiciones moderadas de radiaciones ultravioleta y que había una estirpe mutante de la bacteria particularmente sensible o vulnerable a tales radiaciones. ¿Por qué era así? Posiblemente la bacteria normal tuviera un enzima o sistema de enzimas capaces de reparar las lesiones causadas al ADN por las radiaciones ultravioleta, incrementando la tasa de supervivencia, mientras que la bacteria sensible habría perdido dicha capacidad. Ya que los enzimas, como las demás proteínas, están determinadas por genes, el autor, junto con sus colegas de la Facultad de Medicina de la Universidad de Yale, realizó en 1961 varios experimentos genéticos cuya finalidad era encontrar los genes responsables de la síntesis de enzimas reparadores. Se aprovechó el hecho de que las

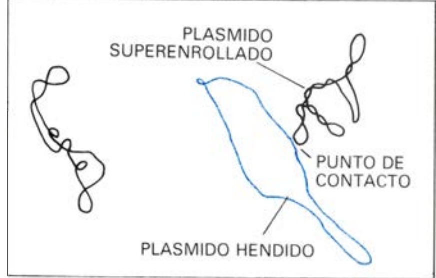
células de *E. coli* intercambian genes mediante conjugación, una forma primitiva de cruzamiento entre células macho y hembra. Se utilizó la estirpe convencional de laboratorio *E. coli* K12 porque era relativamente resistente a las radiaciones y porque se conocía bastante bien su mapa genético: se sabía la localización de muchos genes en la molécula de ADN bicatenario circular que constituye el cromosoma bacteriano.

En la búsqueda de mutantes de *E. coli* sensibles a las radiaciones ultravioleta, a los rayos X, o a ambos, se halló que dicha sensibilidad era una propiedad estable y heredable, que podía transmitirse de unas células a otras vía cruzamientos. Con métodos genéticos se identificaron los genes mutantes y se cartografiaron en el cromosoma de *E. coli*. Algunos de los mutantes sensibles a las radiaciones ultravioleta tenían mutaciones en genes que denominamos *uvrA*, *uvrB* y *uvrC*; otros mutantes, sensibles tanto a las radiaciones ultravioleta como a los rayos X, llevaban mutaciones en un gen que se llamó *lexA*. Se comprobó que los genes *uvr* determinaban la síntesis de enzimas reparadores, mientras que el gen *lexA* determinaba un tipo diferente de proteína, cuya naturaleza se habría de conocer mucho más tarde.

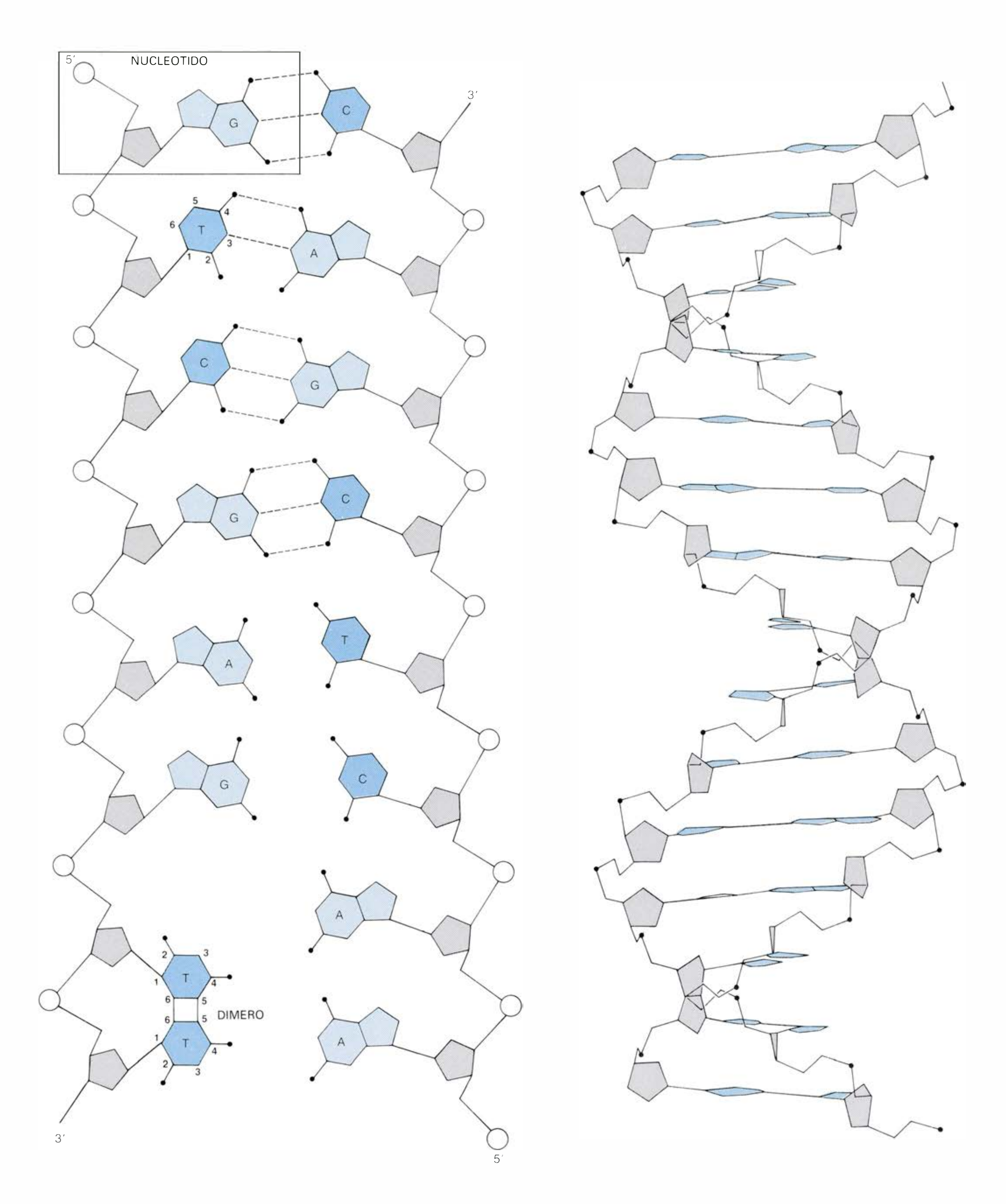
¿En qué mecanismo bioquímico se hallaban implicados los genes *uvr*? Richard B. Setlow y William L. Carrier, del Laboratorio Nacional Oak Ridge, y Richard P. Boyce, Lee Theriot y el autor, en Yale, estudiaron la aparición de dímeros de pirimidina en el ADN de *E. coli* normal y en mutantes sensibles a ultravioleta que eran *uvr*$^-$, es decir, deficientes en alguno de los genes *uvr*. A tal fin, se expusieron las células a radiaciones ultravioleta y se buscaron dímeros de pirimidina en sus respectivos ADN. Los dímeros, como se esperaba, persistían en el ADN cromosómico intacto de las células *uvr*$^-$, pero no en el ADN cromosómico de las bacterias resistentes. Sin embargo, los dímeros no habían desaparecido de las células resistentes. Se encontraban en segmentos de ADN unicatenario de sólo unas pocas bases de longitud. Se acababa de descubrir la "reparación por excisión". El ADN unicatenario que llevaba los dímeros había sido eliminado, presumiblemente, por la acción de uno o más enzimas, determinados por uno o varios genes *uvr*. Supusimos que la reparación del ADN debía completarse con la acción de los enzimas polimerasa y ligasa del ADN, los cuales reemplazarían los nucleótidos eliminados y unirían de nuevo el esqueleto de azúcar-fosfato, respectivamente; la cadena complementaria intacta serviría de molde. Al poco tiempo, David E. Pettijohn y Philip C. Hanawalt, de la Universidad de Stanford, ofrecieron pruebas directas de este tipo de síntesis reparadora.

La reparación por excisión resultó ser uno de los sistemas fundamentales de corrección del ADN, no sólo en bacterias sino también en células de mamíferos (hombre incluido), pues además de eliminar dímeros de pirimidina actúa sobre otros tipos de lesiones del





RECOMBINACION mediada por la proteína RecA, base de uno de los mecanismos de reparación del ADN. La micrografía electrónica, realizada por Era Cassuto y Stephen C. West, en el laboratorio del autor, muestra una de las etapas del proceso de recombinación. Como se indica en el dibujo de la micrografía (*izquierda*), dos copias de un plásmido de ADN (pequeña molécula circular bicatenaria) se aparean en un "punto de contacto"; es ésta una de las etapas iniciales de la recombinación. Uno de los plásmidos es una molécula circular intacta de ADN bicatenario, como demuestra su forma "superenrollada". El plásmido con el que se aparea es idéntico a él, pero en éste se ha eliminado enzimáticamente un pequeño segmento de una de sus cadenas. El hueco resultante permite que el plásmido se "relaje". Sólo se han observado parejas de este tipo al incubar plásmidos intactos con otros similares, aunque con un hueco de cadena sencilla, en presencia de RecA y ATP (la principal fuente de energía de la célula). Las parejas siempre incluían, al menos, una molécula relajada. Cabe presumir que RecA se une a la región de cadena sencilla opuesta al hueco y la alinea con una región homóloga de una cadena del plásmido intacto.



MOLECULA DE ADN (*izquierda*), formada por dos cadenas de nucleótidos. Cada uno de éstos contiene un azúcar (*gris*), un grupo fosfato (*círculos vacíos*) y una base nitrogenada (*color*). Las bases, que pueden ser purinas (adenina, *A*, y guanina, *G*) o pirimidinas (timina, *T*, y citosina, *C*), sobresalen del esqueleto de azúcar-fosfato, lo que posibilita la unión de las cadenas de nucleótidos por puentes de hidrógeno (*líneas a trazos*) entre bases complementarias: *A* sólo puede aparearse con *T* y *G* sólo con *C*. La información genética está impresa en la secuencia de bases. Las dos cadenas tienen polaridades opuestas; la secuencia de bases se lee, convencionalmente, en el sentido de 5' a 3'. En el dibujo, la doble hélice se muestra desenrollada; el plano de las bases, que en realidad es perpendicular al eje de la hélice, se ha girado 90 grados para mostrar la estructura química. Los dímeros de pirimidina, las lesiones más estudiadas en las investigaciones sobre la reparación del ADN, se originan por exposición a la radiación ultravioleta. Se forma un dímero (*abajo*) cuando los dobles enlaces entre los átomos de carbono 5' y 6' de dos pirimidinas adyacentes se abren y constituyen un anillo de cuatro carbonos, aproximando las dos bases y rompiendo sus puentes de hidrógeno con las bases complementarias de la cadena opuesta; la consiguiente distorsión del esqueleto impide la formación de puentes de hidrógeno en dos bases a cada lado del dímero. Los dímeros más frecuentes se forman entre timinas adyacentes de una misma cadena, pero también se pueden formar entre dos citosinas o entre una timina y una citosina. Un modelo esquemático del ADN (*derecha*) muestra la disposición de las dos cadenas de azúcar-fosfato en la estructura de doble hélice. Los detalles del apareamiento de las bases nitrogenadas se aprecian con menor claridad.

ADN. Boyce y el autor vieron al poco tiempo que este sistema reparaba, por ejemplo, las lesiones causadas por la mitomicina *C*, que induce la "soldadura" de las dos cadenas de la doble hélice. Que la mitomicina *C* fuera una droga anticancerígena (porque provoca lesiones en el ADN) y también carcinogénica en células animales (por idéntica razón) era un indicio de que la información que pudiese obtenerse de los estudios de reparación del ADN en bacterias gozaría de amplias aplicaciones.

El descubrimiento de una segunda forma importante de reparación del ADN, muy diferente de la reparación por excisión, ocurrió por vía indirecta. En 1965, A. John Clark y Anne D. Margulies, de la Universidad de California en Berkeley, realizaban estudios de recombinación genética, proceso que permite reunir genes pertenecientes a cromosomas distintos, como acontece, por ejemplo, durante la conjugación bacteriana. Clark aisló mutantes de *E. coli* cuyo ADN no recombinaba en la forma esperada después de la conjugación, y atribuyó la falta de recombinación a una mutación en un gen que llamó *recA*. Al ensayar la resistencia de las células *recA*$^-$ a las radiaciones, se encontró con la sorpresa de que eran extremadamente sensibles a las lesiones causadas por las radiaciones ultravioleta y los rayos X.

W. Dean Rupp y el autor quedaron perplejos ante la aparente relación entre una deficiencia en la recombinación y la sensibilidad a la radiación. Quizá significara que el producto de *recA* tenía algo que ver con un proceso de reparación todavía desconocido. Se comparó la sensibilidad a los ultravioleta de células *uvrA*$^-$ o *recA*$^-$ con la de los dobles mutantes, células con defectos en ambos genes. Los dobles mutantes resultaron ser 50 veces más sensibles que las células *uvrA*$^-$ o las *recA*$^-$, lo que constituía una señal clara de que estaban implicados dos sistemas de reparación muy diferentes. Para simplificar la situación estudiamos la síntesis de ADN en bacterias *uvrA*$^-$, que no podrían realizar la reparación por excisión, aunque conservasen presumiblemente el sistema basado en el producto del gen *recA*, cualquiera que fuese. En 1968 descubrimos la reparación post-replicativa, que actúa sólo después de la replicación del ADN dañado por la radiación ultravioleta.

En una célula que ha estado expuesta a radiación ultravioleta, los segmentos de las cadenas parentales de ADN que no han sufrido alteración se replican con normalidad. Cuando la horquilla de replicación (el punto donde las dos cadenas parentales están desenrolladas y expuestas a los enzimas reparadores) llega a un dímero de pirimidina que no se ha eliminado, el apareamiento de las bases de la cadena dañada resulta imposible. Se interrumpe entonces la síntesis de una cadena nueva, y se reinicia en algún punto de la región no dañada del ADN que sirve de molde, más allá del dímero. Aparecen así "huecos post-replicativos" en las cadenas hijas. El sistema de reparación por excisión no puede actuar sobre estos huecos, ya que requiere una cadena complementaria intacta; la reparación debe basarse en un mecanismo diferente. La secuencia de bases, efectivamente, se pierde en las dos cadenas de una de las dos moléculas hijas del ADN bicatenario formadas en la replicación: una cadena tiene un hueco y la otra un dímero. Hay que buscar algún modo de recuperar la información que falta, si se quiere reparar correctamente el hueco. Nos dimos cuenta de que la secuencia de bases necesaria para ello se hallaba muy cerca, en la otra doble cadena hija formada en la misma horquilla de replicación. ¿Podría un proceso de recombinación sustituir las bases que faltaban en el hueco post-replicativo por una de las cadenas de la molécula bicatenaria "hermana", y de esta forma conseguir una cadena de ADN que tuviera, frente al dímero, la secuencia de bases correcta?

Para detectar esta reparación por recombinación, Rupp y el autor marcaron el ADN con isótopos que permitían distinguir las cadenas parentales de las hijas en razón de su diferente densidad. Pudimos contar el número de huecos post-replicativos y el número de intercambios entre cadenas hermanas. Había, aproximadamente, el mismo número de intercambios que de huecos. La existencia de reparación por recombinación post-replicativa se confirmó al descubrir Ann K. Ganesan, de Stanford, que los dímeros de pirimidina no se localizaban exclusivamente en las cadenas irradiadas con ultravioleta, sino que se distribuían tanto por las cadenas "nuevas" como por las "viejas", a medida que se replicaba el ADN.

La reparación post-replicativa complementa perfectamente a la reparación por excisión. La reparación por recombinación "rellena" los huecos originados, en las cadenas hijas, por los dímeros (u otro tipo de daño) que alcanza la horquilla de replicación antes de haber sido eliminados. La reparación por excisión elimina el dímero (u otro daño) que queda en la cadena parental.

El estudio de la inducción lisogénica, un fenómeno aparentemente no relacionado con la reparación post-replicativa, aportó un mayor conocimiento del papel de *recA*. Se dice que una bacteria es lisogénica cuando, al ser infectada por un bacteriófago (un virus bacteriano), el ADN de éste se integra en el cromosoma bacteriano como profago, en vez de multiplicarse. (La multiplicación conduce a la formación de nuevas partículas de fagos y mata la célula.) El fago lambda es uno de esos virus "temperados", que infecta estirpes lisogénicas de *E. coli*. El ADN del fago lambda puede permanecer "inactivo" en la bacteria, replicándose con el cromosoma bacteriano por muchas generaciones. Cuando la bacteria se expone a radiación ultravioleta o algún otro agente que lesione el ADN, el profago "se induce". Su ADN deja el cromosoma, se replica y dirige la síntesis de proteínas víricas; se producen cientos de nuevas partículas de virus, y la célula estalla, liberándolos. En 1967, Clark, que había descubierto *recA*, junto con Israel Hertman y Salvador E. Luria, del Instituto de Tecnología de Massachusetts, observaron que las células lisogénicas *recA*$^-$ no se inducían, es decir, lambda permanecía como profago inactivo aunque la célula se expusiera a los ultravioleta. En otras palabras, se demostró que el gen *recA* era necesario para la inducción lisogénica.

No es este el único proceso en el que está implicado el gen *recA*. A mediados de la década de los setenta se llegó al conocimiento de que los mismos acontecimientos que precipitaban la inducción lisogénica en una célula irradiada inducían la síntesis de enzimas reparadores del ADN; además, interrumpían la replicación del ADN y la división celular, de tal modo que la célula se alargaba en vez de dividirse y formar dos células hijas; también se incrementaba la tasa de mutagénesis y se estimulaba la síntesis, en grandes cantidades, de cierta proteína, por aquel entonces conocida por proteína *X*. Todas estas respuestas (y otras que no citamos) parecían ser reacciones a una situación de emergencia; Miroslav Radman, de la Universidad Libre de Bruselas, les dio el nombre colectivo de "respuesta *SOS*".

En 1975, Jacqueline George y Marc Castellazzi, del Instituto para la Investigación en Biología Molecular de París, advirtieron que cierta mutación *recA* actuaba a modo de interruptor que desencadenaba la respuesta *SOS*, incluso en ausencia de radiación ultravioleta. Trabajaron con una cepa lisogénica de *E. coli*, portadora de una mutación sensible a la temperatura,

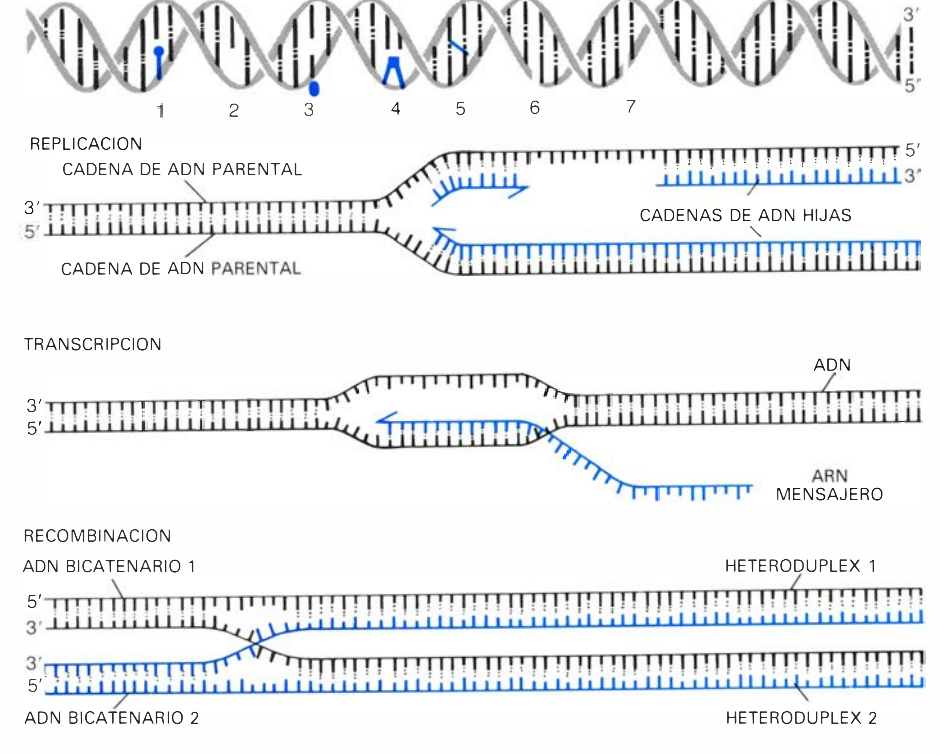
recA441, que era letal para la bacteria si se alcanzaban los 42 grados Celsius. Descubrieron que la muerte celular podía evitarse mediante una segunda mutación en uno cualquiera de los genes *sulA* o *sulB*. Habiendo logrado que las células *recA441* sobreviviesen a 42 grados Celsius, George y Castellazzi demostraron que todas las funciones *SOS* se disparaban a esa temperatura, aun en ausencia de radiación ultravioleta. Si una mutación en *recA* podía desencadenar la respuesta *SOS*, dicha respuesta debía de depender (al menos en parte) del gen *recA*.

Casi al mismo tiempo, Jeffrey W. Roberts, de la Universidad de Cornell, investigaba el mecanismo responsable de la inducción del profago por radiación ultravioleta. Se sabía que el profago lambda permanecía inactivo por la acción de una proteína reguladora llamada represor, que impedía la transcripción de todos los genes del fago (excepto la del que determinaba la síntesis del represor). El represor se adosaba a la región del "operador" en el ADN del fago, impidiendo la unión de la polimerasa del ARN, que de otro modo podría transcribir los genes del fago en ARN. En la búsqueda del mecanismo por el cual la radiación ultravioleta eliminaba al represor de lambda, Roberts y sus colaboradores descubrieron que, durante la inducción, el represor desaparecía. El represor se desprendía del ADN del profago porque se cortaba en dos. Los investigadores de Cornell identificaron y purificaron el enzima responsable del corte: una proteasa (enzima que corta proteínas) con un peso molecular de 38.000. Después de irradiar, detectaban cantidades elevadas de la proteasa en las células de tipo silvestre, pero no así en los mutantes *recA*$^-$. ¿Acaso era la proteasa el producto del gen *recA*? (En las células *lexA*$^-$ tampoco se detectaron cantidades elevadas del enzima.)

Roberts y sus colaboradores encontraron que la proteasa que purificaban a partir de mutantes *recA441* (aquellos en los que las funciones *SOS* están siempre conectadas a 42 grados) era más activa que la proteasa de células de tipo silvestre; por otra parte, un mutante que no llevaba a cabo la inducción del profago, ni siquiera tras la irradiación, carecía de la actividad necesaria para cortar el represor. Análisis posteriores demostraron que esa proteasa era algo poco común. Actuaba muy lentamente, sus blancos específicos eran represores (como el represor de lambda) y su actividad dependía de la presencia de ATP (fuente principal de energía bioquímica) y de ADN unicatenario. Aunque Roberts sospechaba que su proteasa era en efecto el producto del gen *recA*, la confirmación de que no se trataba sólo de un componente minoritario de su preparación no podía obtenerse hasta alcanzar un grado más elevado de purificación.

La identificación final del producto del gen *recA* y su purificación en cantidades interesantes supuso un avance importante. Además de Roberts y sus colaboradores, lo logró un numeroso grupo de investigadores: Peter T. Emmerson y Stephen C. West, en la Universidad de Newcastle upon Tyne, Lorraine J. Gudas y David W. Mount, de la Universidad de Arizona, John W. Little y D. G. Kleid, en el Instituto de Investigación de Stanford, Kevin McEntee, de la Universidad de Chicago, y Hideyuki Ogawa y Tomoko Ogawa, de la Universidad de Osaka. En 1977 se demostró que la proteína *X*, observada por primera vez, en 1971, en células irradiadas por Masayori Inouye y Arthur B. Pardee, de la Universidad de Princeton, y que Gudas halló que faltaba o estaba alterada en los mutantes *lexA*$^-$ y *recA*$^-$, era en efecto la proteína RecA. Se abría con ello el camino a los estudios bioquímicos de la proteína y del mecanismo de la reparación post-replicativa.

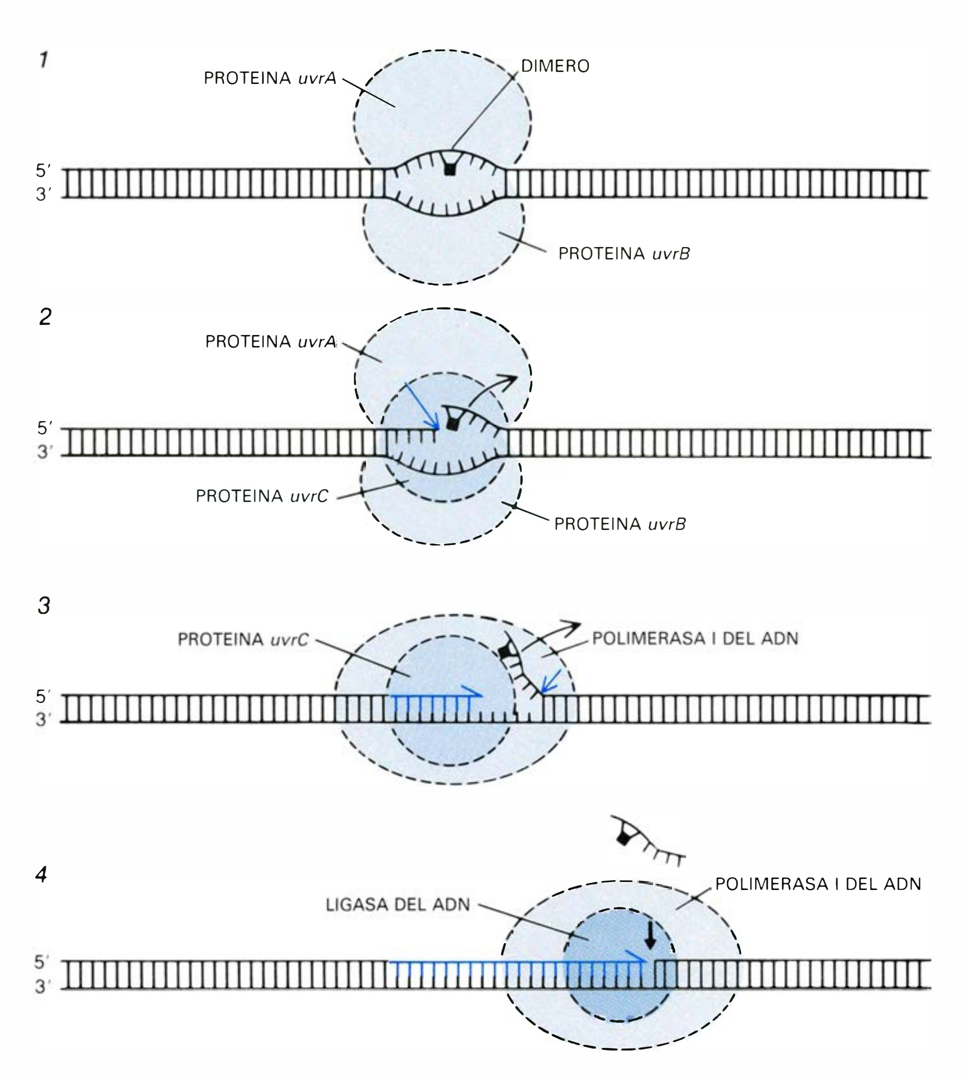
Mientras tanto, seguían investigándose otros aspectos de lo que empezaba a conocerse como "fenómeno *rec-lex*". En 1973, Mount, en Arizona, y K. Brooks Low, en Yale, estudiaron un diploide parcial de *E. coli*: células con un pequeño segmento extra del ADN cromosómico, lo que determina que algunos genes estén presentes en dos versiones. Observaron que los mutantes *lexA*$^-$ eran genéticamente dominantes, es decir, una célula portadora de un gen *lexA* silvestre y de un gen *lexA*$^-$ mutante expresaba las características de la célula mutante, incluida la sensibilidad a los daños inducidos por rayos X. La dominancia en un gen mutante suele indicar que el gen tiene algo que ver con la regulación genética. Ciertas mutaciones en el gen *lacI* de *E. coli*, gen que determina la síntesis del represor que controla la expresión de los genes involucrados en el metabolismo de la lactosa, son dominantes. Cabía, pues, que el gen *lexA* determinase una proteína represora que regulase ciertos enzimas de la reparación del ADN.



VARIOS TIPOS DE LESIONES DEL ADN. Como se indica en el esquema de la doble hélice (*arriba*), las bases pueden alterarse químicamente por alquilación, hidroxilación, etcétera (*1*), o pueden incluso eliminarse (*2*). Unas veces se forma un complejo químico por adición de una molécula voluminosa (*3*), otras, se forman dímeros de pirimidina (*4*), se "sueldan" químicamente las cadenas (*5*) o se adhiere una proteína a una de las cadenas. Puede romperse una cadena (*6*) o ambas (*7*). La mayoría de estas lesiones alteran el apareamiento de las bases o la continuidad del molde, e interfieren el funcionamiento normal del ADN. La replicación y la transcripción dependen del apareamiento de las bases para la síntesis de dos nuevas cadenas hijas de ADN, o de una cadena de ARN mensajero; también la mayoría de los procesos de recombinación dependen del apareamiento para la alineación correcta de las regiones homólogas de ADN.

Mount razonó que los mutantes *lexA*⁻ eran sensibles a la radiación porque fabricaban un represor defectuoso, que no se liberaba después de la irradiación. (Ello explicaría la dominancia de *lexA*⁻. Aun cuando hubiera en la célula moléculas normales de represor, que se liberasen del ADN, la proteína *lexA*⁻ defectuosa continuaría uniéndose y reprimiento los enzimas reparadores.) A partir de una estirpe portadora de una mutación *lexA*⁻, denominada *lexA3*, Mount buscó células que fueran de nuevo resistentes a la luz ultravioleta. Encontró una que aún llevaba la mutación original y, además, una nueva mutación designada *lexA51*, que situó en el locus *lexA*, o muy próxima a él. Aparentemente, esta mutación afectaba a la función *lexA* de tal forma que los genes antes reprimidos por la proteína de *lexA* podían ahora expresarse, se sintetizaban los enzimas reparadores y se restañaban los daños producidos por la radiación ultravioleta. Se seguía de ello que los mutantes *lexA51* fabricaban un represor incapaz de unirse correctamente al ADN; por consiguiente, la síntesis de enzimas reparadores y todas las demás funciones *SOS*, sin excluir el aumento de la síntesis de RecA, estarían siempre "conectadas", incluso en ausencia de radiación ultravioleta u otro agente lesivo para el ADN. Se demostró así que la respuesta *SOS* no dependía sólo de *recA*, sino también de *lexA*, con una alta probabilidad de que el producto de este último gen fuese una proteína represora que controlase todos los genes implicados en la respuesta *SOS*.

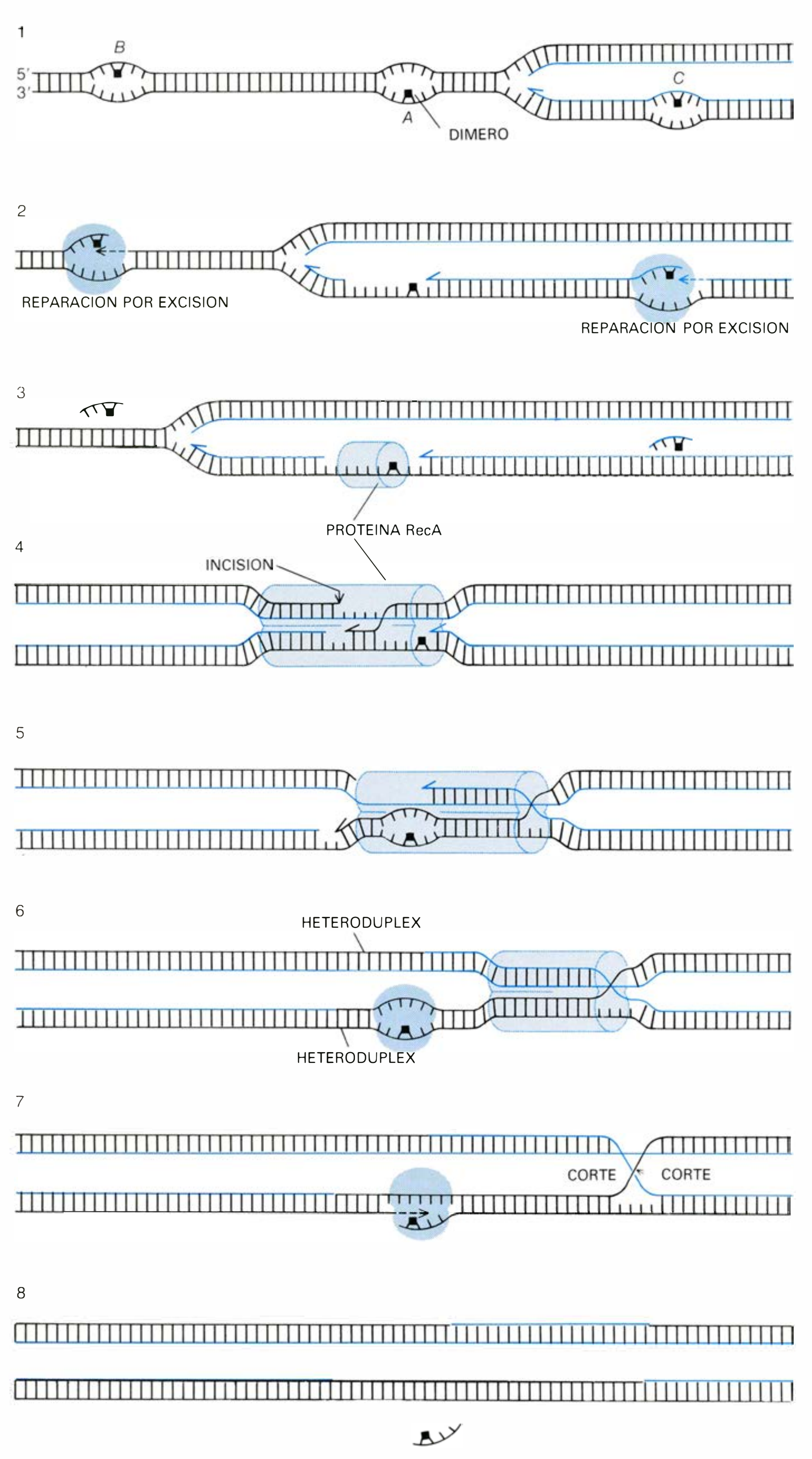
El paso siguiente consistía en identificar la proteína codificada por *lexA*. Little (por entonces en Arizona) y Mount lo consiguieron, utilizando para ello técnicas de ADN recombinante. Aislaron el gen *lexA*, lo unieron a un plásmido (pequeño trozo circular de ADN cromosómico que se replica autónomamente en la célula bacteriana), insertaron el plásmido recombinante en *E. coli* y lograron identificar y purificar la proteína: la así llamada LexA tenía un peso molecular de 24.000. Para confirmar el hipotético papel represor de LexA la sometieron a la acción de la proteasa de *recA*. Su descubrimiento fue trascendental: la proteína RecA rompía LexA. Además, esta rotura ocurría con más rapidez que la del represor de lambda, lo que vino a sugerir que LexA era el sustrato primario de la proteasa. Mark Ptashne y Roger Brent, de la Universidad de Harvard, demostraron que la proteína LexA reprimía al gen *lexA*: el gen sufría represión por su propio producto, como ocurre con el gen del represor de lambda y con algunos otros genes.



LA REPARACION POR EXCISION elimina los dímeros de pirimidina y restaña otras lesiones del ADN en un proceso de "corte y remiendo". En este proceso intervienen tres enzimas especiales, determinados por los genes *uvrA*, *uvrB* y *uvrC*, junto con la polimerasa I del ADN y la ligasa del ADN. Primero, las proteínas *uvrA* y *uvrB* se unen al sitio dañado (*1*). A continuación, y quizá con la ayuda de la proteína *uvrC*, realizan una incisión (*flecha de color*) en el extremo 5' de la región dañada (*2*). En presencia de la proteína *uvrC*, la polimerasa del ADN se coloca en el lugar de la incisión y añade nucleótidos al extremo 3', siguiendo las reglas del apareamiento de bases. La polimerasa realiza una segunda incisión en la cadena para eliminar la región dañada (*3*); la polimerasa continúa "digiriendo" nucleótidos y reemplazándolos al mismo tiempo por otros nuevos, desplazándose la incisión un poco más a la derecha (*4*). La ligasa del ADN termina la reparación, uniendo la cadena en el lugar de la incisión (*flecha gruesa*).

Si LexA es un represor, ¿a qué otros genes inhibe, además del propio? Cynthia J. Kenyon y Graham C. Walker, del Instituto de Tecnología de Massachusetts, idearon un ingenioso experimento para determinar qué genes de *E. coli* se expresaban en respuesta a un daño inferido a su ADN. Aplicaron un método desarrollado por Malcolm J. Casadaban y Stanley N. Cohen, de la Facultad de Medicina de la Universidad de Stanford, para insertar aleatoriamente el gen *lacZ* en el cromosoma de una estirpe que carecía de todos los genes que controlan el metabolismo de la lactosa. Después, buscaron células que sólo metabolizaran la lactosa en presencia de un agente capaz de atentar contra el ADN (así la mitomicina *C*). En dichas células, pensaron, algunos de los genes que normalmente se encontraban reprimidos se expresarían ahora en respuesta a una lesión del ADN. Sin embargo, la función normal del gen se había perturbado por la inserción de *lacZ*, y era *lacZ* lo que se expresaba. Cuando se identificaron los genes que se activaban en estas condiciones (por la pérdida de su función normal), resultaron ser los genes *uvrA* y *uvrB*, genes que determinan la síntesis de enzimas reparadores, y el gen *umuC*, relacionado con el aumento de la mutagénesis que sigue a la irradiación con ultravioleta.

El experimento había permitido identificar varios genes aparentemente reprimidos por la proteína LexA. Barry



REPARACION POST-REPLICATIVA. Actúa sobre un dímero de pirimidina (*A*) que bloquea la replicación; por ésta se entiende el proceso a lo largo del cual las cadenas parentales (*negras*) se desenrollan en la horquilla de replicación y se sintetizan dos cadenas hijas (*color*) (*1*); otros dímeros (*B* y *C*) se eliminan mediante la reparación por excisión. El dímero *A* impide el apareamiento de bases en un segmento de una de las cadenas parentales, produciendo un hueco post-replicativo frente a una región de ADN de cadena sencilla (*2*). La proteína RecA (*color claro*) se une a la región de cadena sencilla (*3*) y sitúa frente a ella la región homóloga de la molécula bicatenaria hermana. Cuando se consigue el apareamiento homólogo, un enzima corta la molécula bicatenaria (*4*). La proteína RecA enfrenta el extremo libre de la cadena parental de la molécula bicatenaria con el hueco, produciéndose un intercambio de cadenas (*5*). La polimerasa del ADN puede ahora reparar el heterodúplex superior. Con la secuencia correcta situada en frente del dímero *A* y con la proteína RecA liberada, los enzimas de la reparación por excisión pueden entonces eliminar el dímero *A* (*6*). Finalmente, un enzima practica dos cortes en el lugar del intercambio de cadenas (*7*), terminando el proceso de recombinación con un saldo de dos moléculas heterodúplex enteras.

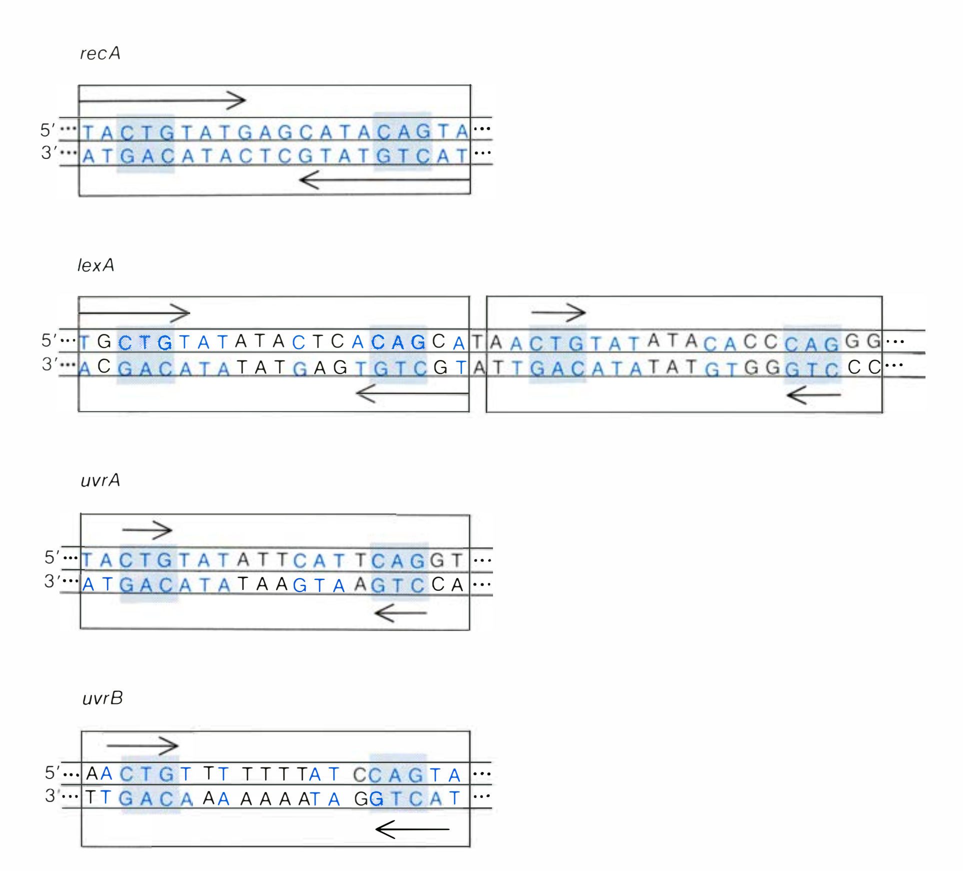
M. Kacinski, Aziz Sancar y Rupp emprendieron la búsqueda de pruebas directas de esta regulación. Encontraron que el mutante *lexA51* de Mount, en el que las funciones *SOS* se encontraban siempre activadas, pues no había represión por LexA, exhibía mayor actividad *uvrA* y *uvrB* sobre el ADN dañado que las células normales. El incremento de actividad estimulaba la reparación del ADN: Jeffrey Auerbach, trabajando en el laboratorio del autor, detectó una mayor supervivencia en los mutantes *lexA51* que en las células de tipo silvestre, debido posiblemente a una mayor actividad *uvrA* y *uvrB* en la estirpe mutante.

Si la proteína LexA es realmente un represor de un grupo de genes implicados en la reparación del ADN debe unirse a regiones operadoras cercanas al comienzo de cada uno de estos genes, por lo que cabría esperar que dichas regiones presentasen secuencias de nucleótidos similares. La secuencia nucleotídica del gen *recA* ha sido determinada por Toshihiro Horii y los Ogawa, en Osaka, y por Sancar y Rupp; la secuencia de la región operadora de *lexA*, por el grupo de Osaka y por Little y Mount, y por Sancar y Rupp, parte de los genes *uvrA* y *uvrB*. En la región operadora de estos genes hallamos una secuencia muy similar, de unos 20 nucleótidos; esta "caja *SOS*" tiene (en diferente grado, según los genes) la forma de secuencias simétricas invertidas, característica de los sitios de unión del represor. En particular, el gen *lexA* posee dos secuencias adyacentes, casi idénticas, análogas a los múltiples sitios de unión presentes en el operador del fago lambda. Tanto Little y Mount como Brent y Ptashne han purificado la proteína LexA y han demostrado que se une específicamente a las cajas *SOS* de *recA* y *lexA*.

Podemos ya describir, en líneas generales, la serie de fenómenos interrelacionados que se dan cita en la respuesta *SOS* y en su regulación. En una célula en crecimiento, que no ha sufrido daño alguno, los genes reparadores del ADN se encuentran casi totalmente inhibidos por LexA; ello no obsta para que se sintetice suficiente cantidad de represor destinado a ocupar las regiones operadoras, y suficiente proteína *uvr* que lleve a cabo esporádicas reparaciones por excisión. Cuando las células se exponen a una dosis significativa de radiación ultravioleta, los dímeros de pirimidina que no pueden reparar las pocas moléculas de proteína *uvr* presentes en el estado no inducido originan huecos post-replicativos. Las proteínas RecA que en aquel momento

haya en la célula se unen al ADN unicatenario que queda frente a los huecos; el ADN estimula la actividad proteasa de RecA y se induce la respuesta *SOS*. La proteína RecA destruye a los represores LexA, levantando la represión a *recA* y a los otros genes regulados de la misma forma; se sintetizan grandes cantidades de proteína RecA, que se unen al ADN unicatenario todavía libre y, posteriormente, a regiones bicatenarias contiguas. En estas condiciones, la proteína LexA se destruye en cuanto se sintetiza y los genes regulados permanecen activos mientras haya ADN unicatenario en la célula para estimular la actividad proteasa de RecA. Cuando la reparación post-replicativa termina y todo el ADN se encuentra en forma bicatenaria, RecA deja de actuar como proteasa. La proteína LexA recién sintetizada no se destruye y puede cumplir su función represora; la célula vuelve al estado no inducido.

La respuesta *SOS* estimula la reparación eficaz del ADN, pues aumenta la disponibilidad de enzimas para la reparación por excisión y abre la vía de la reparación post-replicativa. Los enzimas determinados por *uvrA, uvrB* y *uvrC* actúan conjuntamente en la reparación por excisión, reconociendo las bases dañadas e iniciando el proceso de excisión. Puesto que las células sobreviven a miles de lesiones de este tipo, y aunque el proceso sea lento, las células deben de disponer de grandes cantidades de enzimas. Rupp y sus colaboradores han demostrado que, normalmente, una célula no dañada tiene sólo 10 o 20 moléculas del enzima determinado por *uvrA*; el aumento de la síntesis que sigue a la inducción *SOS* permite que se disponga de suficiente enzima para ablacionar eficazmente un gran número de dímeros. Se está estudiando actualmente en varios laboratorios el mecanismo exacto de acción de los tres enzimas, y para incrementar el nivel de su síntesis se han clonado los genes que determinan los enzimas en plásmidos presentes en las células con un alto número de copias. Además, el método de la "maxicélula", desarrollado por Sancar y Rupp, permite identificar y purificar enzimas, aun cuando no se disponga de ensayos funcionales para detectar las proteínas.



REGIONES OPERADORAS próximas a los cuatro genes controlados por la proteína represora LexA. Tienen una secuencia similar, de aproximadamente 20 pares de bases; estas secuencias comunes son los sitios de unión del represor. (Hasta ahora, la unión sólo se ha demostrado en los casos de *recA* y *lexA*.) Se señalan en color los pares de bases de los operadores *lexA*, *uvrA* y *uvrB* que son idénticos a los del operador *recA*; los pares de bases que difieren aparecen en negro. Todos los sitios de unión, llamados "cajas *SOS*", incluyen secuencias repetidas e invertidas, o palindrómicas, (*flechas*), que se leen idénticamente en el sentido 5' → 3' de ambas cadenas. Una característica común es la presencia de las secuencias *CTG* y *CAG* (*fondo coloreado*) separadas por 10 bases. El gen *lexA* tiene dos cajas *SOS* casi idénticas.

El lector habrá notado que RecA desempeña dos funciones muy diferentes: actúa de proteasa en su papel regulador y es un enzima de recombinación, capaz de manipular cadenas de ADN. Su primera función en la recombinación, como ocurre en la reparación post-replicativa, consiste en estimular el apareamiento homólogo entre dos moléculas de ADN, es decir, alinear de forma adecuada las regiones con secuencias de bases complementarias para permitir el intercambio de cadenas. En 1979, McEntee, en Chicago, George M. Weinstock e I. Robert Lehman, en la Facultad de Medicina de Stanford, y Takahiko Shibata, Chanchal DasGupta, Richard P. Cunningham y Charles M. Radding, en Yale, demostraron que la proteína RecA podía inducir apareamientos homólogos entre un ADN unicatenario y otro bicatenario. El grupo de Radding y Era Cassuto y West, en el laboratorio del autor, encontraron que la proteína inducía apareamientos homólogos entre dos moléculas bicatenarias circulares, siempre que una de ellas portara una región unicatenaria equivalente a un hueco post-replicativo, pero no lo inducía en el caso de que ambas moléculas bicatenarias estuvieran intactas.

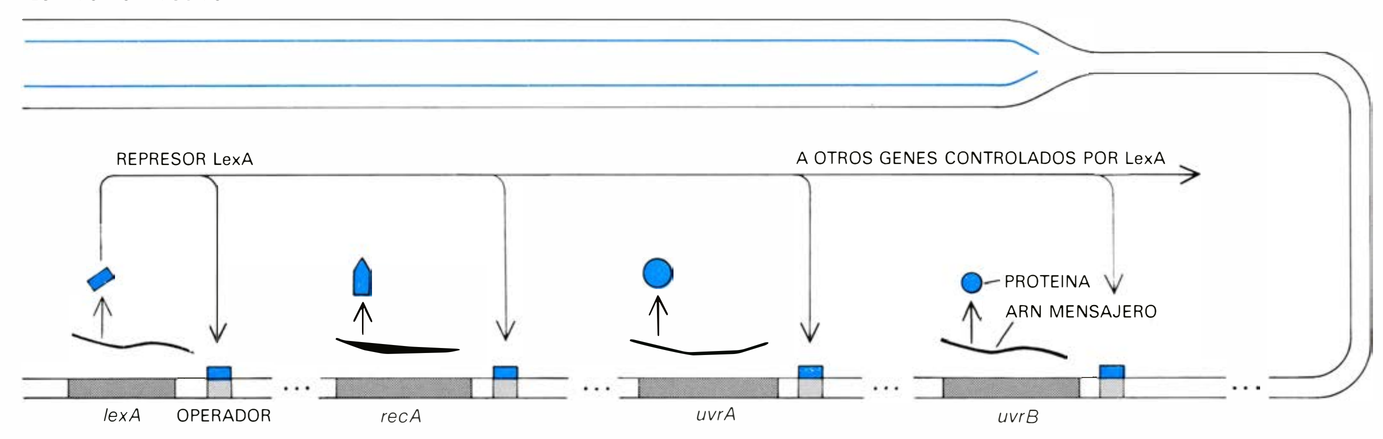
No se conoce aún el mecanismo exacto del apareamiento; parece probable, sin embargo, que la molécula bicatenaria con un hueco unicatenario se alinee de forma inespecífica con la molécula que permanece entera y, luego, por mediación de la proteína RecA, se desplacen las moléculas hasta enfrentar a las bases complementarias. El estudio del requerimiento energético de este proceso aporta un poco de luz a la cuestión. El ATP constituye el suministrador verosímil de energía. Se observó que las regiones unicatenarias y las bicatenarias con un hueco de cadena sencilla estimulaban la hidrólisis del ATP (para recabar la energía necesaria) mediada por RecA; mas las moléculas bicatenarias intactas no ejercían tal estímulo. Evidentemente, la proteína adoptaba una configuración activa en las moléculas bicatenarias con huecos de cadena sencilla. El apareamiento homólogo necesita grandes cantidades de proteína RecA (una molécula por cada cinco pares de bases); éstas confieren a la cadena que presenta el hueco una estructura fibrosa de soporte que permite poner en contacto a las moléculas de ADN, de forma que se muevan una sobre otra hasta establecer los contactos complementarios. Este movimiento recuerda al de una oruga: el ADN que lleva los huecos se alarga y se contrae a medida que la proteína RecA se "pega" y se "despega", degradando al propio tiempo ATP.

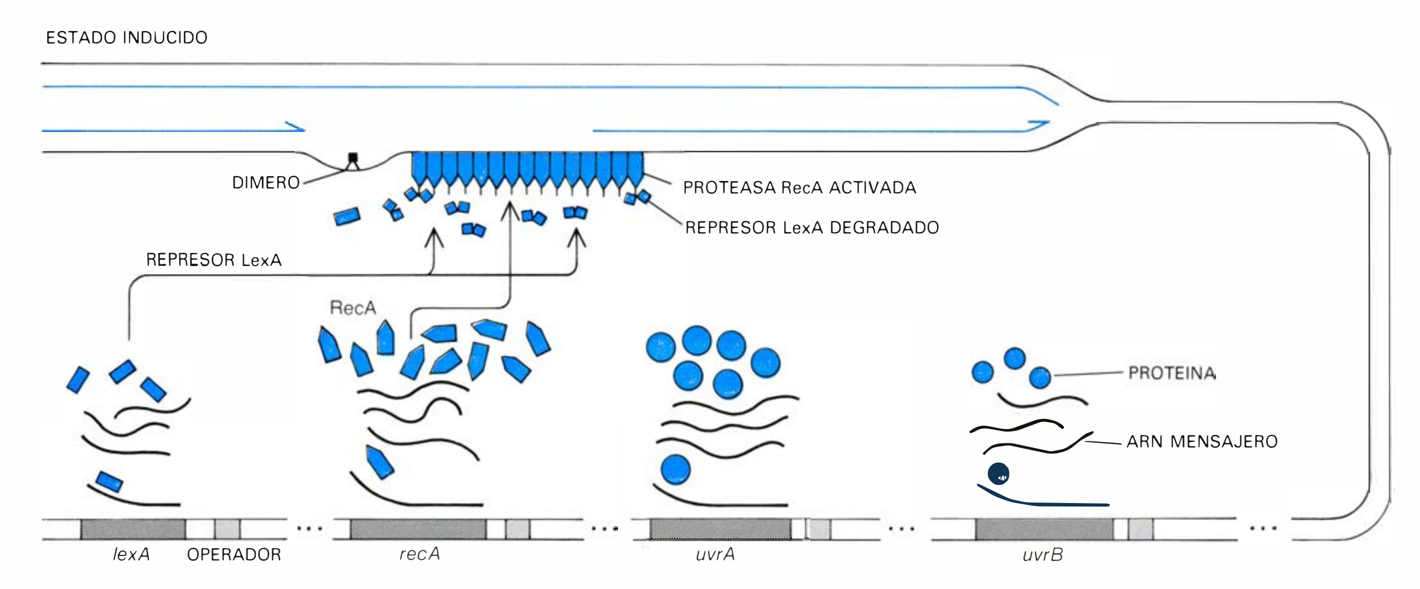
Una vez que las moléculas de ADN

establecen contactos relativos a la complementariedad en una región determinada, RecA promueve el intercambio de cadenas. Utilizando moléculas bicatenarias con las regiones unicatenarias adecuadas, West ha demostrado que RecA realiza un intercambio recíproco entre dos regiones bicatenarias; resulta de ello la formación de dos heterodúplex en los que las cuatro cadenas tienen las parejas cambiadas. El proceso discurre mediante un cruzamiento entre cadenas sencillas, hecho éste que fue propuesto en 1965, como una estructura intermedia en la recombinación, por Robin Holliday, del Instituto Nacional de Investigaciones Médicas de Londres. En presencia de ATP, RecA dirige el intercambio de cadenas a medida que el lugar donde ha ocurrido el entrecruzamiento avanza a lo largo de la molécula, proceso llamado "movimiento de ramas". Para completar la reparación se necesitan otros enzimas, además de RecA: una nucleasa que corte las cadenas en diversas etapas del intercambio, una polimerasa del ADN que "rellene" los huecos con nuevas bases, una ligasa para unir el esqueleto de azúcar-fosfato, y quizás algún otro enzima más.

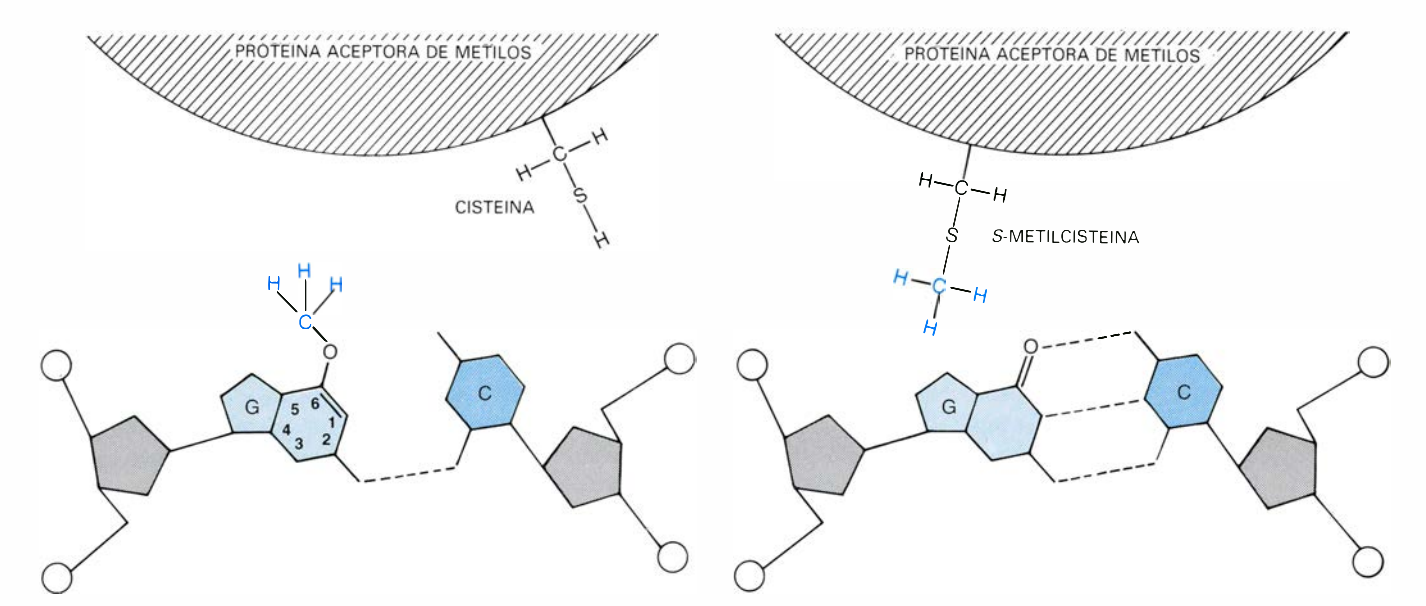
Con anterioridad se ha mencionado que el aumento de la mutagénesis parece ser uno de los aspectos de la respuesta *SOS*. Hace algunos años, Jean J. Weigle, del Instituto de Tecnología de California, estudiando la mutagénesis inducida por los ultravioleta en el fago lambda, descubrió que la frecuencia de mutación y la supervivencia de los fagos irradiados aumentaban cuando éstos infectaban a células que habían sido a su vez irradiadas unos 30 minutos antes. Evelyn M. Witkin, de la Universidad Rutgers, y Jun-ichi Tomizawa, que trabaja actualmente en el Instituto Nacional de Artritis, Metabolismo y Enfermedades Digestivas, encontraron que no se daba inducción de mutaciones por radiación en las bacterias *recA*$^-$ o *lexA*$^-$.

Witkin ha sugerido que un tipo especial de reparación "con tendencia al error" sería el responsable del incremento en la supervivencia y la mutagénesis en las bacterias de tipo silvestre irradiadas. Radman y ella propusieron que la respuesta *SOS* induciría la aparición de una polimerasa del ADN especial, particularmente tolerante con las mutaciones presentes en el ADN "molde" y capaz de replicar ADN, observando las reglas de apareamiento de bases de una forma menos estricta que la polimerasa normal. Una polimerasa con tendencia al error aumentaría el ni-





EL SISTEMA REGULADOR basado en *lexA* y *recA* está "en reposo" durante el crecimiento normal en ausencia de daños al ADN (*arriba*). El represor LexA, unido a los operadores de *lexA*, *recA*, *uvrA* y *uvrB* y algunos otros genes más, mantiene la síntesis del ARN mensajero y proteínas al bajo nivel característico de las células no inducidas. Un daño en el ADN, suficiente para producir un hueco post-replicativo, activa la respuesta *SOS* (*abajo*). La proteína RecA se une al ADN unicatenario que queda frente al hueco, se activa su actividad proteasa y degrada al represor LexA. Libres del represor funcional, los genes controlados por LexA se "conectan" y aumenta la velocidad de síntesis de proteínas. (Aunque se sintetiza más LexA, éste se degrada mientras la proteína RecA esté activada por la presencia de ADN unicatenario.) Terminada la reparación del ADN, RecA no degrada ya al represor; la proteína LexA recién sintetizada se une otra vez a los operadores y la célula vuelve al estado no inducido. (Los dibujos de este artículo son de Patricia J. Wynne.)



LA METILACION de una guanina por el agente químico nitrosoguanidina (*MNNG*) se repara mediante un nuevo proceso, que pone de manifiesto una respuesta adaptativa a la lesión del ADN. La nitrosoguanidina coloca grupos metilo (CH_3) en diversos puntos del ADN, por ejemplo, en el átomo de oxígeno que está en la posición número 6 de la guanina (*izquierda*). La metilación provoca la rotura de los puentes de hidrógeno que unen a la guanina con la citosina. La reparación corre a cargo de una proteína que se sintetiza, por adaptación, cuando la célula se expone a una concentración baja de nitrosoguanidina. Una cisteína de la proteína actúa como aceptor de grupos metilos: incorpora el grupo metilo y devuelve la guanina a su estado original (*derecha*).

vel de supervivencia a expensas de una mayor frecuencia de mutación. Stuart M. Linn, David Lackey y Sharon W. Krauss, de Berkeley, han detectado cantidades muy pequeñas de una polimerasa del ADN con tendencia al error en células donde se ha inducido un aumento en la síntesis de RecA. La mutagénesis, en particular por ultravioleta, comprende una serie de complejos mecanismos que sólo ahora empiezan a ponerse de manifiesto.

La reparación por excisión y la reparación post-replicativa parecen ser dos mecanismos fundamentales y necesarios para restañar los daños del ADN, pero hay mecanismos más específicos cuando se trata de lesiones circunscritas. Así, bases individuales incorrectas o alteradas se eliminan mediante ciertos enzimas, las glicosilasas, muy específicas para determinados tipos de bases dañadas.

De particular interés resulta un sistema de reparación descubierto recientemente. Radica su atracción en su mecanismo, poco usual, y en el hecho de que parece representar una segunda clase de respuesta inducida por las lesiones genéticas, bastante distinta de la respuesta *SOS*. Se llama respuesta adaptativa; aligera la gravedad del daño producido al ADN por la metilación (adición de un grupo CH_3) de una base guanina.

La respuesta adaptativa fue descubierta en 1977 por Leona Samson y John Cairns, a la sazón en los laboratorios de la Fundación Imperial de Investigación sobre el Cáncer, en Londres. Investigaban los posibles efectos de una exposición prolongada a bajas concentraciones de sustancias químicas potencialmente carcinogénicas. Estudiaron el efecto, en *E. coli*, de exposiciones prolongadas a bajas concentraciones de un potente mutágeno llamado *N*-metil-*N*'-nitro-nitrosoguanidina (nitrosoguanidina o *MNNG*), de uso común en el laboratorio. Sorprendentemente, bastaban dos horas de incubación con sólo un microgramo de nitrosoguanidina por cada mililitro de medio de cultivo para que las bacterias adquiriesen una fuerte resistencia a un tratamiento posterior con una concentración de la droga cientos de veces más elevada. Esta respuesta adaptativa, que se atribuyó a un gen denominado *ada*, duraba una o dos horas. A lo largo de ese intervalo, el nivel de mutagénesis y de letalidad celular que resultaba de la metilación era mucho más bajo que en células no adaptadas. En las células adaptadas, la sensibilidad a los daños inducidos por radiación ultravioleta no se veía afectada, por lo que la respuesta adaptativa no podía considerarse mera peculiaridad de la respuesta *SOS*.

Thomas Lindahl y sus colaboradores, en la Universidad de Göteborg, analizaron los mecanismos bioquímicos responsables de la respuesta adaptativa. Aunque la nitrosoguanidina metila las bases en varias posiciones, la respuesta adaptativa parece inducirse cuando se ve afectado el átomo de oxígeno designado por O^6; el grupo metilo insertado sobresale de lo que se conoce como "surco mayor" de la doble hélice de ADN. Mientras que la mayoría de las bases alteradas se eliminaban del ADN por acción de enzimas reparadores adecuados, que las reemplazaban, Lindahl encontró que en este caso la guanina permanecía en su lugar y que era el grupo metilo lo que se eliminaba. El trabajo corría a cargo de una proteína especial, sintetizada en abundancia sólo durante la respuesta adaptativa, que actuaba como aceptor de metilos. Una característica del proceso era la limitación de su capacidad de acción. Cuando un aminoácido (la cisteína) de una molécula de proteína reparadora aceptaba un grupo metilo (formándose *S*-metilcisteína), la capacidad reparadora de la molécula se agotaba.

Quedan aún por determinar todas las ramificaciones de la respuesta adaptativa. Muchos aspectos de la enzimología de la respuesta *SOS* y los detalles de su control por LexA son hoy objeto de análisis. Sin duda han de descubrirse más sistemas reparadores inducidos. El estudio de estos sistemas en bacterias, que aglutina la labor de muchos laboratorios, está aún lejos de haberse cerrado. La propia investigación científica está empeñada, en no pocas de sus líneas de trabajo, en el desentrañamiento de los enzimas reparadores de ADN y de la respuesta a los daños genéticos en organismos superiores, mamíferos incluidos. La razón de ese afán es obvia: los indicios cada vez más firmes de que los daños inferidos al ADN y su reparación son factores a tener en cuenta en el cáncer humano y quizás en otras enfermedades, así como en el proceso de envejecimiento.

Ciencia y sociedad

Un calendario popular y la astronomía española en el siglo X

Un programa de investigación de una cierta amplitud proporciona a veces resultados inesperados que podrían calificarse de "sub-productos" del mismo a no ser por su obvio interés. Este es el caso que presento rápidamente aquí: un grupo de trabajo barcelonés ha iniciado un programa de estudio sobre "La astronomía española en el siglo XIII", uno de cuyos capítulos se proponía revisar una serie de manuscritos de cómputo españoles de este siglo, con el fin de demostrar la coexistencia de dos corrientes astronómicas en la España de la época: una, mejor conocida, de origen árabe y representada fundamentalmente por la obra de Alfonso X, y una segunda, de raíces latinas, que sobreviviría en los tratados de cómputo eclesiástico. Con este propósito emprendí el estudio, junto con mi compañero José Martínez Gázquez de la Universidad Autónoma de Barcelona, del *Liber Regius* (manuscrito 167) del Museo Episcopal de Vich y ambos descubrimos, con sorpresa, que el mencionado libro contenía una nueva traducción latina del *Calendario de Córdoba* del siglo X. Esta nueva traducción parece haber sido realizada en Cataluña y, posiblemente, entre 1228 y 1235.

El *Calendario de Córdoba* es un calendario popular que contiene rúbricas generales para cada uno de los meses del año solar, así como indicaciones particulares para ciertos días de cada mes. Estas indicaciones son de índole muy diversa: astronómicas, meteorológicas, agrícolas, dietéticas, festividades cristianas y musulmanas. En cierto sentido es un precedente de los calendarios populares actuales como *El Zaragozano* y el *Calendari del Pagès*. Sus autores fueron un médico musulmán, Arib b. Sad, y un obispo mozárabe, Rabí b. Zayd o Recemundo, quienes lo dedicaron al califa cordobés al-Hakam II (961-976). Este hecho explica el que en el texto sobreviva una doble tradición: árabe por una parte (tanto preislámica como helenizada) y latino-cristiana por otra. El texto árabe del *Calendario* nos ha sido conservado en un único manuscrito, muy defectuoso, escrito en caracteres hebreos al que se añadía una traducción latina, del siglo XII, atribuida a Gerardo de Cremona. Ambos textos han sido editados más de una vez y la mejor edición, hasta el momento, es la del profesor Charles Pellat (Leiden, 1961). La pobreza de la tradición textual explica el interés que puede presentar la nueva traducción latina conservada en el *Liber Regius* de Vich, ya que este texto puede servirnos para entender mejor ciertos pasajes del original.

Desde el punto de vista astronómico, el *Calendario de Córdoba* tiene el interés de ser testimonio de la difusión en la España Musulmana de dos series de tablas astronómicas orientales: por una parte el *Sind-Hind*, que remonta a una obra india del siglo VII, introducida en Bagdad hacia el 771, y en Córdoba alrededor del 850; por otra, las tablas de al-Battani, astrónomo sirio que vivió entre el 858 y el 929. Ahora bien, hasta el momento presente nadie parece haberse preocupado por el interés que tienen, en nuestro *Calendario*, ciertos valores numéricos que aparecen, aproximadamente, dos veces por mes. Se trata de valores correspondientes a las alturas meridianas del Sol, longitud de la sombra correspondiente a dicha altura, duración del día y de la noche y, finalmente, duración de la aurora y del crepúsculo. La presencia de estas series de valores en el *Liber Regius*, con la correspondiente aparición de variantes, me ha movido a estudiar los procedimientos utilizados para calcularlos como único método seguro para deslindar los valores originales de aquellos que proceden de un error de copista o de traductor. Presentaré, aquí, brevemente, un avance de los resultados de este trabajo.

1. *Alturas meridianas del Sol*: el *Calendario* nos da estas alturas veintitrés veces a lo largo del año. Si utilizamos la fórmula elemental:

$$h_m = (90^\circ - \varphi) \pm \delta$$

(para h_m = altura meridiana del Sol; φ = latitud del lugar; δ = declinación del Sol en el día de referencia), podremos fácilmente sacar algunas consecuencias de los valores suministrados por el *Calendario* para las alturas meridianas correspondientes a los equinoccios (52° 30'), solsticio de verano (76° 20') y solsticio de invierno (28° 40'). Efectivamente, estos valores se ajustan bien a una latitud de 37° 30' y a una declinación máxima (oblicuidad de la eclíptica) de 23° 50'. En lo que respecta a la latitud, el parámetro utilizado aquí no es muy corriente para Córdoba (el más usual en la Edad Media es 38° 30'), pero se encuentra documentado en uno de los manuscritos de las *Tablas de Toledo*. En cuanto a la oblicuidad de 23° 50', es obvio que se trata del resultado de redondear el parámetro 23° 51' utilizado por Ptolomeo. Por otra parte, partiendo de los parámetros anteriores, es sencillísimo reconstruir los valores de la declinación solar correspondientes a cada altura meridiana, valores que me han servido para recomputar las cifras que proporciona el *Calendario* para la duración del día, por una parte, y del crepúsculo por otra. Como consecuencia de lo apuntado en este apartado, parece poco probable que los valores de la altura meridiana del Sol que nos da el texto cordobés se basaran en la observación. Es más fácil, en cambio, que se calcularan a partir de una tabla de declinación solar por más que, en el momento presente, resulte difícil precisar exactamente cuál es la fuente utilizada.

2. *Sombras correspondientes a las alturas meridianas del Sol*. Empecemos, aquí, por aclarar qué se entiende por "sombra" en astronomía medieval: sea AB [*véase el dibujo de la página 39*] un *gnomon* perpendicular al suelo HC. El Sol se encuentra en S y el rayo de luz SAC forma un triángulo rectángulo ABC con el *gnomon* AB y el suelo HC. En este triángulo el ángulo C mide la altura del Sol sobre el horizonte (arco SH). Evidentemente:

$$\text{cotg } C = \frac{BC}{AB}$$

Por tanto, la cotangente de la altura del Sol será igual a la sombra BC dividida por la altura del *gnomon* AB. Esta última será siempre una constante y la tradición hispánica suele considerar que el *gnomon* mide 12 dígitos, de lo que se desprende que el valor de la *sombra* medieval suele ser doce veces el de la cotangente de unas tablas trigonométricas actuales.

Ahora bien, el análisis de los valores de las sombras tal como aparecen en nuestro *Calendario* nos reserva algunas sorpresas. En primer lugar, el *gnomon* utilizado es la estatura de un hombre erecto, lo que implica que los valores de la cotangente son fundamentalmente los mismos que podemos obtener hoy con unas tablas trigonométricas calculadas para un radio (o *gnomon*) igual a la unidad. Se trata del primer caso conocido en España de funciones trigonométricas computadas para $g = 1$, por más que no deba darse excesiva importancia al hecho, ya que la fi-

nalidad perseguida no es la de simplificar el cálculo trigonométrico (algo que están buscando, en el Oriente musulmán, un grupo de matemáticos y astrónomos contemporáneos, aproximadamente, del *Calendario de Córdoba*). Más adelante, en el siglo XI, Ibn Muadh de Jaén computará la primera tabla de tangentes hispánica conocida utilizando $g = 1$.

El *Calendario* nos da 25 valores de la sombra. Si multiplicamos por 12 cada uno de estos 25 valores, podremos comprobar que, en 19 casos, obtendremos números enteros. Esto hace pensar que las sombras del *Calendario* proceden de una tabla de sombras computada para $g = 12$, tal como puede verse en la primera tabla de la derecha, donde aportamos una muestra de media docena de estos valores. No puedo aún identificar la fuente de la que derivan pero sí sugerir que se trata, posiblemente, de la yuxtaposición de dos tablas de sombras distintas. Una de ellas correspondería al paso del Sol por los puntos medios de cada uno de los signos zodiacales y tendría la siguiente estructura (I = enero, II = febrero, etcétera):

I	II	III	IV	V	VI
XII	XI	X	IX	VIII	VII
21	16	11	7	4,5	3

La segunda, en cambio, daría las alturas meridianas correspondientes a la entrada del Sol en cada signo zodiacal:

	I	II	III	IV	V	VI
XII	XI	X	IX	VIII	VII	
22	19	14	9	6	4	3

Se trata de un tipo de tablas, de estructura muy primitiva, bien conocido en la Antigüedad Clásica y que se están poniendo de relieve, recientemente, en la Edad Media Islámica por los trabajos de D. A. King.

3. *Duración del día*: aquí los resultados obtenidos son menos espectaculares. Los valores que da el *Calendario* son, en general, bastante correctos, tal como puede comprobarse con los que he recalculado, utilizando una latitud de 37° 30' y los valores de la declinación derivados de la tabla de alturas meridianas, mediante la fórmula:

$$\text{Duración del día} = \frac{[90^\circ + sen^{-1}(tg\delta\ tg\varphi)]\ 2}{15}$$

[*véase la segunda tabla de esta página*]. Esta fórmula era conocida en la Córdoba del siglo x ya que aparece en las tablas de al-Battani que, como hemos visto, es una de las fuentes astronómicas citadas explícitamente por el *Calendario*.

4. *Duración del crepúsculo*: los astrónomos musulmanes entienden por crepúsculo el tiempo que el Sol tarda entre su paso por un círculo de altura negativo determinado (los más corrientes son los que corresponden a las alturas − 17°, − 18° y − 19°) y su aparición por el horizonte (crepúsculo matutino o aurora), o bien entre su paso por el horizonte y el momento en que alcanza el mencionado círculo de altura negativo (crepúsculo vespertino). Evidentemente, se trata de un caso particular de un problema mucho más general omitido por la tradición helenística, pero que atrajo mucho la atención de los astrónomos indios y musulmanes: establecer la hora en función de la altura del Sol. Los autores del *Calendario* podrían haber utilizado la fórmula, enunciada por al-Battani, así como también por Alfonso X en su *Libro de las Taulas*, que nos da el ángulo horario (t) en función del arco semidiurno (D), de la altura del Sol para un momento determinado (h) y de la altura meridiana del Sol (h_m); t valdría:

$$\frac{D \pm vers^{-1}[vers\ D(1 - sen\ h/sen\ h_m)]}{15}$$

en la que *vers* es una función trigonométrica poco usual hoy en día, el seno verso, que se define de la siguiente manera:

$$vers\ A = 1 - cos\ A$$

Es obvio que la duración de la aurora y del crepúsculo no ha sido computada, en nuestro *Calendario*, de acuerdo con este procedimiento, ya que, para una altura de − 18°, la diferencia entre los valores del *Calendario* y los que se obtienen con esta fórmula llega a ser de unos 30 minutos para el solsticio de invierno. Ahora bien, al comparar la duración del día en el *Calendario* con la del crepúsculo, mi compañero Manuel García Doncel me hizo observar que la duración del crepúsculo es, aproximadamente, la mitad de la duración del día (o sea el arco semidiurno expresado en horas) dividida por 4. Teniendo presente esta observación, revisé los pro-

FECHA	ALTURA (Cal. Córdoba)	SOMBRA (Cal. Córdoba) (1)	COTG (valor actual) (2)	DIFERENCIA (1) − (2)	SOMBRA x 12
2-I	29° 40'	1 3/4	1,7556	− 0,0056	21
18-III	52° 30'	3/4	0,7673	− 0,0173	9
16-VI	76° 20'	1/4	0,2432	0.0068	3
16-VIII	65° 15'	1/2	0,4610	0,0390	6
15-IX	54° 30'	2/3	0,7133	− 0,0466	8
16-XII	28° 40'	1 5/6	1,8291	0,0042	22

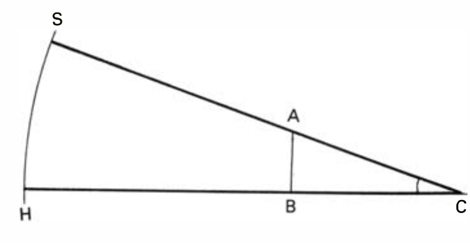
Valores de las sombras proyectadas por el Sol a su paso por el meridiano

FECHA	DECLINACION	DURACION DIA (Cal. Córdoba) (1)	DURACION DIA (recomputada) (2)	DIFERENCIA (1) − (2)
1-I	− 22° 50'	9 h. 30 m.	9 h. 29 m. 12 s.	48 s.
16-III	0°	12 h.	12 h.	0
16-VI	23° 50'	14 h. 40 m.	14 h. 38 m. 31 s.	1 m. 31 s.
16-VIII	12° 45'	13 h. 20 m.	13 h. 19 m. 59 s.	1 s.
15-IX	2°	12 h. 12 m.	12 h. 12 m. 17 s.	− 17 s.
15-XII	− 23° 50'	9 h. 20 m.	9 h. 21 m. 29 s.	− 1 m. 29 s.

Duración del día

FECHA	DURACION CREPUSCULO (Cal. Córdoba) (1)	DURACION CREPUSCULO (recalculada para $h = -17°$) (2)	DIFERENCIA (1) − (2)
1-I	1 h. 8 m. 34 s.	1 h. 6 m. 44 s.	1 m. 50 s.
19-III	1 h. 24 m.	1 h. 24 m. 18 s.	− 18 s.
1-VI	2 h.	1 h. 41 m. 52 s.	18 m. 8 s.
1-VIII	1 h. 55 m.	1 h. 37 m. 11 s.	17 m. 49 s.
1-IX	1 h. 32 m.	1 h. 28 m. 59 s.	3 m. 1 s.
15-XII	1 h. 9 m.	1 h. 5 m. 34 s.	3 m. 26 s.

Duración del crepúsculo matutino y vespertino



La sombra como función trigonométrica

cedimientos arcaicos, pre-trigonométricos, conocidos para deducir la hora en función de la altura del Sol, y hallé un método, establecido por el astrónomo indio del siglo VII, Brahmagupta, y conocido en el mundo árabe al menos desde fines del siglo x, según el cual:

$$t = \frac{D}{cotg\ h + 1}$$

(siendo, como antes, t = ángulo horario, D = arco semidiurno y h = altura negativa del Sol en un momento dado). Resulta obvio que, de acuerdo con la observación de García Doncel:

$$cotg\ 17^o + 1 = 4{,}2709$$
$$cotg\ 18^o + 1 = 4{,}0777$$
$$cotg\ 19^o + 1 = 3{,}9042$$

Utilizando este viejo procedimiento de origen indio he recalculado los valores del *Calendario de Córdoba* [*véase la tabla tercera de la página anterior*] para $h = -17^o$, parámetro con el que se obtienen los mejores resultados. Los resultados obtenidos son aceptables salvo para el período de tiempo en torno al solsticio de verano, en el que las cifras del *Calendario* parecen claramente corruptas. Es posible, por tanto, que un método similar al expuesto haya sido el utilizado para calcular los valores de nuestro texto.

Para terminar me limitaré a decir que el *Calendario de Córdoba* es uno de los textos hispánicos, con materiales astronómicos, más antiguos conocido. Las series de valores numéricos que en él aparecen llaman la atención por la disparidad de niveles de resolución de los problemas planteados: soluciones, probablemente, pre-trigonométricas para los valores de las sombras correspondientes a las alturas meridianas del Sol y para la duración del crepúsculo; plenamente trigonométricas, en cambio, para la duración del día. El conjunto me hace pensar en una astronomía tradicional, muy primitiva, que recoge materiales de fuentes diversas y de valor muy distinto. Nada nos garantiza que estas series de cifras estuvieran calculadas, originariamente, para Córdoba: podrían muy bien derivar de unas tablas computadas para cualquier lugar del mundo islámico cuya latitud fuera de 37° 30'. (Julio Samsó.)

Relevancia de la histona H1

Las moléculas biológicas se asocian entre sí para regular su función o adquirir propiedades nuevas. A este fenómeno asociativo no parece sustraerse la biomolécula "reina", aquella que contiene la información para la síntesis de todas las biomoléculas de cada ser vivo, y, por tanto, para el crecimiento y reproducción de éstos: *el ácido desoxirribonucleico*, abreviadamente ADN. Pues bien, en las células que tienen núcleo el ADN se encuentra dentro de éste unido a una cantidad parecida o superior de proteínas. La mayor parte de dichas proteínas nucleares forman un grupo relativamente constante en los distintos seres vivos, y reciben el nombre de histonas. El resto de proteínas nucleares agrupa un conjunto mucho más heterogéneo y variable, y son denominadas no-histonas.

La estructura y función de las histonas y de las no-histonas, así como de las asociaciones que forman con el ADN, empiezan a ser conocidas, en los últimos años, gracias a los esfuerzos de numerosos investigadores. Pese a ello, persisten todavía grandes lagunas por explicar. Hay, sin embargo, un convencimiento generalizado de que las proteínas nucleares, al unirse al ADN, y en algunos casos al forzarlo a que adopte ciertas disposiciones en el espacio, modulan la transmisión del mensaje genético (facilitándolo o dificultándolo, como mínimo). Así, los cambios de actividad del ADN a lo largo de la vida de las células suelen ir acompañados de cambios en la estructura y empaquetamiento de la cromatina (nombre que se da a las asociaciones de ADN y proteínas nucleares). Esta puede pasar, por ejemplo, de un estado difuso al estado compacto de cromosomas. Dichos cambios pueden correlacionarse con variaciones en la estructura de las proteínas nucleares componentes o en el tipo de proteínas presentes.

La caracterización bioquímica sistemática de las proteínas histonas recibió un gran impulso a partir del año 1965, destacando en dicha tarea los grupos de investigación de Butler, Johns y Bradbury en Inglaterra, Allfrey, Bonner y Cole en Estados Unidos y de Palau y Subirana en España. A partir de núcleos de células, se consiguió aislar cinco tipos principales de histonas, denominadas H1, H2A, H2B, H3 y H4. Al estudiar el tipo de aminoácidos que forman dichas histonas y el encadenamiento secuencial que adoptan dentro de ellas (estructura primaria o secuencia), se apreció que la histona H1 presenta unas características claramente distintas de las demás histonas, como son: un mayor número de aminoácidos componentes (alrededor de 200 frente a 100-150 de las demás), una diferente distribución de dichos aminoácidos en la secuencia y una variabilidad de secuencia mucho mayor entre las histonas H1 de distintos seres vivos.

Sin embargo, han sido los avances realizados en los últimos años sobre la subestructura de la cromatina los que han definido inequívocamente el carácter singular de la histona H1 respecto de las demás histonas. Según ello, dos moléculas de cada una de las cuatro histonas, H2A, H2B, H3 y H4, forman agrupaciones octaméricas sobre las que se enrolla el ADN de forma repetitiva, dando lugar a estructuras arrosariadas. Cada una de dichas agrupaciones de ADN e histonas se denomina nucleosoma. Pues bien, la histona H1 en la cromatina tiene una disposición externa a cada uno de dichos nucleosomas, colocándose en el punto en que el ADN inicia y finaliza al mismo tiempo su enrollamiento sobre las demás histonas. Además, en ciertas condiciones, las cadenas de nucleosomas pueden también enrollarse, dando lugar a superestructuras, probablemente de tipo cilíndrico con un agujero en el centro (solenoide), proceso que viene facilitado por la presencia de la histona H1.

Otra avance significativo en la caracterización de la histona H1 y de proteínas relacionadas, realizado consecutivamente a los anteriores, ha consistido en la resolución parcial de su conformación espacial y de sus propiedades funcionales. A ello han contribuido, de manera importante, el Laboratorio de Biofísica del Politécnico de Portsmouth (Inglaterra) así como el Instituto de Biología Fundamental y el Departamento de Bioquímica de la Universidad Autónoma de Barcelona (España). F. X. Avilés, autor de esta nota, ha participado en dichos trabajos, en uno y otro centro.

Los estudios realizados hasta el presente sobre la histona H1 son un ejemplo de cómo se puede obtener un conocimiento aproximado de la conformación espacial de una proteína no-cristalizable. En la actualidad, la única técnica que permite obtener la estructura tridimensional completa de una proteína es la difracción de rayos X de cristales de ella. En caso de que la proteína no se pueda cristalizar, se procede a su estudio conformacional en disolución mediante la aplicación de técnicas que dan sólo una información parcial de ella: forma, tamaño, disposición externa o interna de ciertos aminoácidos, relaciones de vecindad de algunos aminoácidos, etcétera. Gracias a dichas técnicas se ha demostrado que la histona H1 está conformada por tres regiones: una central, replegada y de carácter globular, que denominaremos G-H1, y dos laterales, desplegadas, que denominaremos N-H1 y C-H1. La región globular comprende unos 80 ami-

noácidos y tiene un diámetro de 2,9 nanómetros. Esta región globular no se encuentra de hecho en el centro de la H1, ya que la región extrema C-H1 es bastante más larga (más del doble) que la otra región extrema N-H1. La histona H1 entera tiene, pues, una forma alargada con un glóbulo en el interior.

También se tienen datos sobre cómo se repliega la cadena proteica de la H1 dentro de su región globular. Se ha calculado que aproximadamente un 30 por ciento de los aminoácidos de dicha región se encuentran formando parte de estructuras helicoidales cortas. Sin embargo, quedan ciertas incógnitas sobre los límites de dichas hélices y sobre la existencia de otros tipos de plegamiento, así como sobre su disposición relativa en el espacio. Se sabe, además, que en la H1 todos los aminoácidos aromáticos (aquellos formados por estructuras cíclicas) se encuentran encerrados dentro de la región globular y se conoce, en algunos casos, cómo establecen contactos entre ellos y con los aminoácidos vecinos. Con la información actual, no se puede aún dibujar el camino espacial que sigue la cadena proteica dentro de la zona globular de la H1, que es uno de los últimos fines de este tipo de estudios.

En toda esta investigación ha sido muy útil el estudio de grandes fragmentos de la histona H1, conteniendo una de las tres regiones citadas o combinaciones de ellas; ello ha permitido delimitar las propiedades conformacionales y funcionales de la histona. Además, la comparación entre histonas H1 distintas, procedentes de varias especies diferentes, ha facilitado la resolución de sus rasgos fundamentales.

¿Qué ,interés tiene el conocimiento de la conformación de la histona H1? Al igual que en muchas otras proteínas, en la resolución de su función biológica interviene ese conocimiento conformacional, que complementa los estudios de carácter claramente funcional. Así, se ha comprobado que entre el punto de entrada y de salida del ADN en el nucleosoma (después de dar dos vueltas sobre el agregado de histonas H2A + H2B + H3 + H4) queda una distancia de unos 2,9 nanómetros. Esta distancia coincide con el diámetro de la región globular de la histona H1, citado antes. Por otro lado, hay pruebas de que el fragmento globular de la histona H1 (el G-H1) se une a los extremos del ADN que sobresalen del nucleosoma, ya que en presencia de este fragmento dichas regiones del ADN quedan protegidas frente a la acción de enzimas rompedores. Todo ello ha permitido lanzar la teoría de que la histona H1 tiene, entre otras, la función de estabilizar el nucleosoma, colocándose en la cromatina "in vivo" entre las zonas de entrada y salida del ADN en cada nucleosoma, uniéndose a ellas por su región globular. Las regiones extremas, desplegadas, de la H1 se dispondrían sobre las regiones del ADN que enlazan nucleosomas.

Otra indicación funcional proviene del carácter desplegado y rico en aminoácidos cargados positivamente de las regiones extremas de la H1. Según ello, dichas regiones serían las principales responsables de la capacidad de la histona H1 para unirse a la cromatina y para empaquetar sus cadenas de nucleosomas. Este proceso da lugar a superestructuras que culminan, en último término, en la formación del cromosoma. Sin embargo, se desconoce el modo concreto por el que la H1 realiza dicha función: si se coloca en el interior o en el exterior de los enrollamientos supuestamente helicoidales de nucleosomas, si las regiones extremas de la H1 yacen a lo largo de las espiras o si se colocan transversalmente, uniendo espiras distintas del enrollamiento.

Cabe añadir que se han encontrado indicios de que en las regiones de la cromatina en las que la histona H1 ha promovido la formación de superestructuras, el ADN subyacente es incapaz de transcribirse o replicarse: es inactivo. Además, basándose en estudios sobre la composición de la H1 a lo largo de la vida de las células, se ha sugerido que ciertas modificaciones químicas de dicha histona son las responsables de las variaciones en el grado de compactación y actividad de la cromatina. Dichas modificaciones químicas producirían cambios en la conformación de la histona H1 y en su unión con el ADN, lo cual se reflejaría en la estructura de la cromatina. La veracidad de esta hipótesis, sin embargo, está aún por comprobar. La resolución de la conformación de la histona H1 y de sus variaciones funcionales tiene, pues, un interés que excede al del conocimiento puramente académico, dada su función aparentemente clave en la vida de la célula. (F. X. Avilés Puigvert.)

El enigma de Cancho Roano

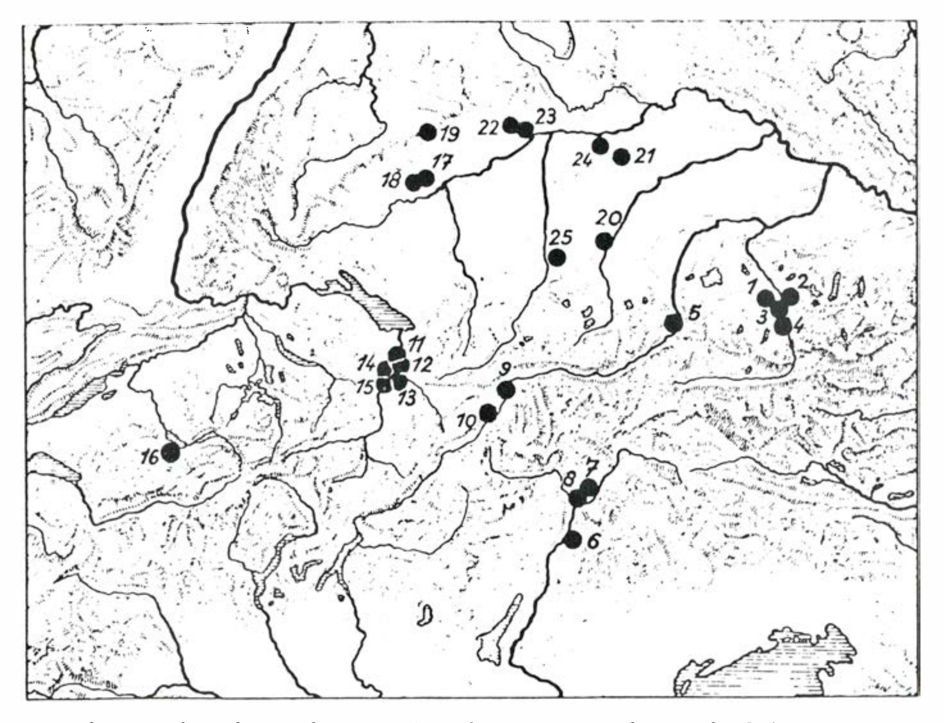
Las excavaciones realizadas por el Instituto de Prehistoria y Arqueología de la Universidad de Barcelona en el término de Zalamea de la Serena (Badajoz) y en el paraje concreto de Cancho Roano han explorado un ingente cúmulo de ceniza, de hasta cuatro metros de espesor, que cubre los restos de un edificio de adobe levantado sobre un basamento de piedras de dos metros de altura y 24 × 24 de superficie [J. Maluquer de Motes y R. Pallarés, *El Palau-Santuari de Zalamea de la Serena, Badajoz (Extremadura)*, Barcelona, Nacrem, 1981]. El monumento ha sido interpretado, y así resalta en el título de la publicación, como un *palacio-santuario* de los siglos VI y V a.C., que en un cierto momento –hacia el año 410 a.C. según análisis del carbono-14– sucumbió a consecuencia de un incendio, y desde entonces, y hasta el año 370 a.C. (por ende, por espacio de unos 40 años) fue lugar de cremación para las gentes de la comarca de la actual Zalamea. Entre las cenizas se encuentran grandes cantidades de huesos de animales domésticos (cápridos, cerdos, algunos équidos); cerámica de muy diversas fábricas, entre las que no falta la griega de Atenas; objetos muy varios de bronce (las patas de la estatua de una cabra, bocados de caballo, una campanilla, amén de muchas piezas de adorno y de aplicación), etcétera.

Tras visitar el lugar de la excavación y leer la memoria de las campañas realizadas por sus directores, nos hemos permitido discrepar de varias de sus interpretaciones, viendo en las ruinas, no un palacio, sino, simplemente, los restos de un edificio sobre el cual no nos atrevemos a pronunciarnos –ni creemos que nadie deba hacerlo sin que la excavación del mismo adelante mucho más, que durante buena parte de su existencia, si no durante toda ella, fue escenario de grandes cremaciones y sacrificios de carácter ritual por los habitantes protohistóricos de la zona, a quienes Plinio llama "célticos" y considera rama del tronco de los lusitanos [A. Blanco Freijeiro, "Cancho Roano, un monumento protohistórico en los confines de la antigua Lusitania", *Boletín de la Real Academia de la Historia*, CLXXVIII (1981) 225-242].

Lo primero que traen a la memoria el enorme depósito de cenizas de Cancho Roano y los muchos despojos de sacrificios cruentos e incruentos que encierra son los datos que el historiador y filósofo estoico Posidonio, que residió una larga temporada en España a comienzos del siglo I a.C. recoge sobre los pobladores de esta zona y que Estrabón reproduce en su *Geografía* (III, 3, 7): "Sacrifican a Ares [el dios de la guerra indígena, al que las inscripciones llaman *Coso*] machos cabríos, prisioneros y caballos, y hacen hecatombes de cada especie, como los griegos, o para decirlo con palabras de Píndaro, *todo lo sacrifican por centenares*".

¿Será por ventura Cancho Roano una ilustración tangible de estos ingentes, abrumadores sacrificios de centenares de víctimas? La diferencia de fechas entre unos restos de los años 550-370 a.C. y unas noticias del siglo I a.C. ha impulsado a los señores Maluquer y Pallarés a pasar por alto estas fuentes, para ir en busca de paralelos arqueológicos en el mundo de los palacios siriohititas de los siglos IX-VIII a.C. y de templos escalonados más antiguos aún: una operación intelectual que nosotros, mostrando sin rebozo nuestra discrepancia, consideramos efecto extemporáneo del *Mirage Orientale* que tanto encandilaba a nuestros abuelos.

Es significativo, a nuestro modo de ver, que siendo Posidonio y Estrabón dos grecoasiáticos, no adviertan la menor similitud entre el ritual lusitano y el de los pueblos orientales antiguos y contemporáneos, entre los que ellos habían nacido y vivido, y apunten en cambio a las hecatombes griegas, citando a Píndaro expresamente, como lo más afín a aquello que tenían delante. Y es que, en efecto, en Grecia hay unas manifestaciones muy semejantes, conocidas no sólo por los textos sino por la arqueología: los "altares de ceniza" del templo de Zeus en Olimpia (un cúmulo de ceniza compacta de seis metros de altura) y del de Hera en Samos, para referirnos sólo a algo muy general y que en Atenas, según Pausanias (V,



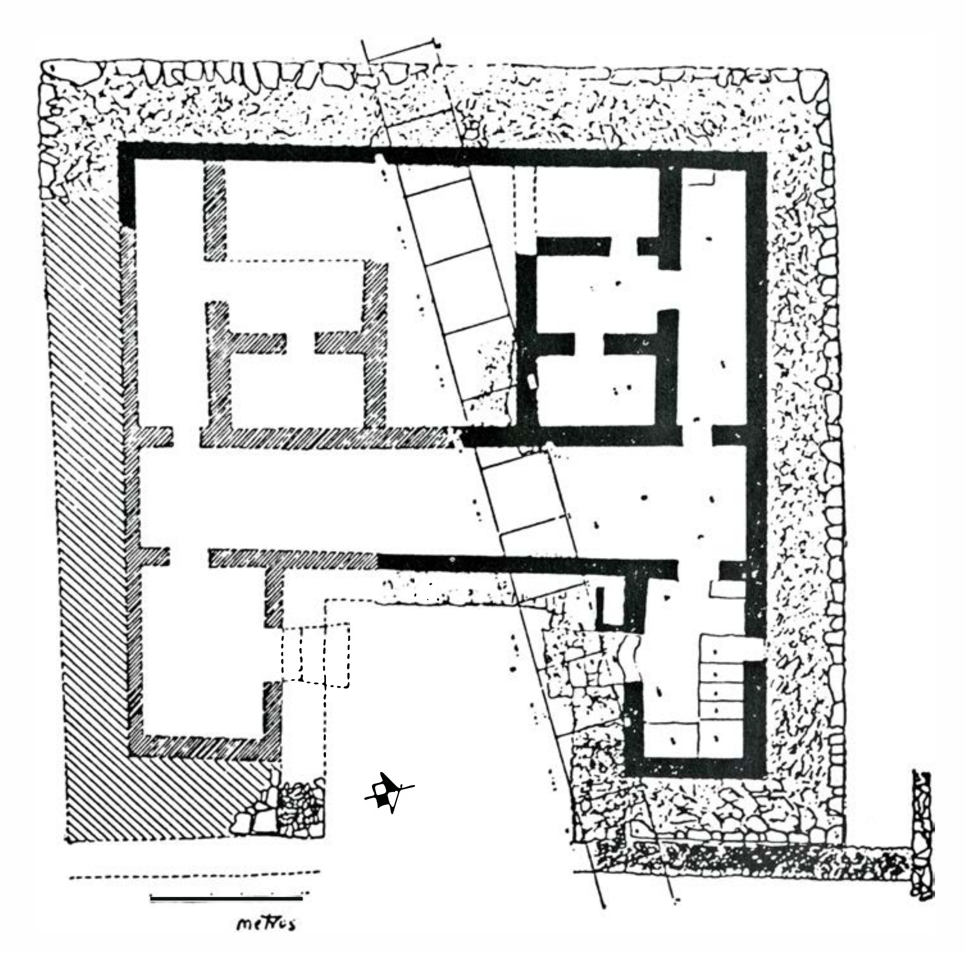
Distribución de "altares de ceniza" prehistóricos en el área de Centroeuropa: 1. Langacker, 2. Goiserberg, 3. Hellbrunner Berg, 4. Dürrnber, 5. Wasserfeldbühel, 6. Schlern, 7. Laugen, 8. Altbrixen, 9. Landeck, 10. Mottata, 11. Feldkirch, 12. Heidenburg, 13. Scheibenstuhl, 14. Schneller, 15. Gutenberg, 16. Eggli, 17. Osterstein, 18. Hägelesberg, 19. Messelstein, 20. Gauting, 21. Ratzenhofen, 22. Weiherberg, 23. Rollenberg, 24. Stätteberg, 25. Auerberg. (Según W. Krämer.)

13,8), se conocía como "altares improvisados".

¿Pero son estos "altares de ceniza" algo privativo de Grecia o parte de un mundo más amplio y al que todo el Oriente es ajeno? Quien interprete a Cancho Roano por esta vía no tardará en averiguar que los "altares de ceniza", o "lugares de holocausto" (*Brandopferplätze*), como los llaman los alemanes, son una institución corrientísima en el círculo alpino y danubiano de época protohistórica, o sea, que son algo archiconocido de los estudiosos de las culturas de los "campos de urnas" y del Hallstatt, en la zona que justamente la arqueología y la filología señalan como posible cuna de gran parte de los indoeuropeos de España, a los que pertenecen lingüísticamente los lusitanos.

La forma en que estos "lugares de holocausto" suelen aparecer es la de un túmulo más o menos alto (4 metros medía originariamente el de Langacker, en la Alta Baviera, que encabeza la serie) y de forma ovalada, con un diámetro de entre 20 y 35 metros. Todos ellos se caracterizan por encerrar cantidades enormes de huesos de animales, quemados o sin quemar (bueyes, ovejas, cabras, cerdos); centenares de vasos cerámicos; objetos pertenecientes al vestido (fíbulas, alfileres) o al adorno corporal (brazaletes), y grandes masas de ceniza y de carbón de leña. En algunos (Langacker), el túmulo se alza sobre una plataforma empedrada a la que se superponen varios pisos de barro (hasta siete), tan quemados por el fuego, que tienen un color rojo de ladrillo; en otros casos aparecen bajo la ceniza muros de mampostería (Feldkirch), o de grandes piedras (Stätteberg), "provistas aquí y allá de estrechos vanos o nichos en los que cabría un hombre en pie". La cerámica permite datar estos yacimientos desde la Edad del Bronce hasta la cristianización; pero ninguno alcanza por sí solo tanta duración; los de vida más larga pueden haber durado un milenio. Lo que sí cabe asegurar es que muchos de ellos se hallaban en su apogeo al mismo tiempo que Cancho Roano.

Parece que los sacrificios no eran todos del mismo tipo. Los huesos sin quemar, probablemente, proceden de banquetes rituales; los quemados, de sacrificios de víctimas. Muchos de estos serían hechos a nombre de una comunidad, como los de Olimpia. Otro indicio que también apunta a Olimpia es el hecho de que varios "lugares de holocausto" no pertenezcan a un poblado, sino que se encuentren en pleno campo, tanto en lo alto de los montes como en lo más hondo de los valles. Además de



Plano del monumento de Cancho Roano (Zalamea, Badajoz), según J. Maluquer y R. Pallarés. La franja de cuadrados que cruza la planta del edificio de este a oeste corresponde a las primeras campañas de excavación.

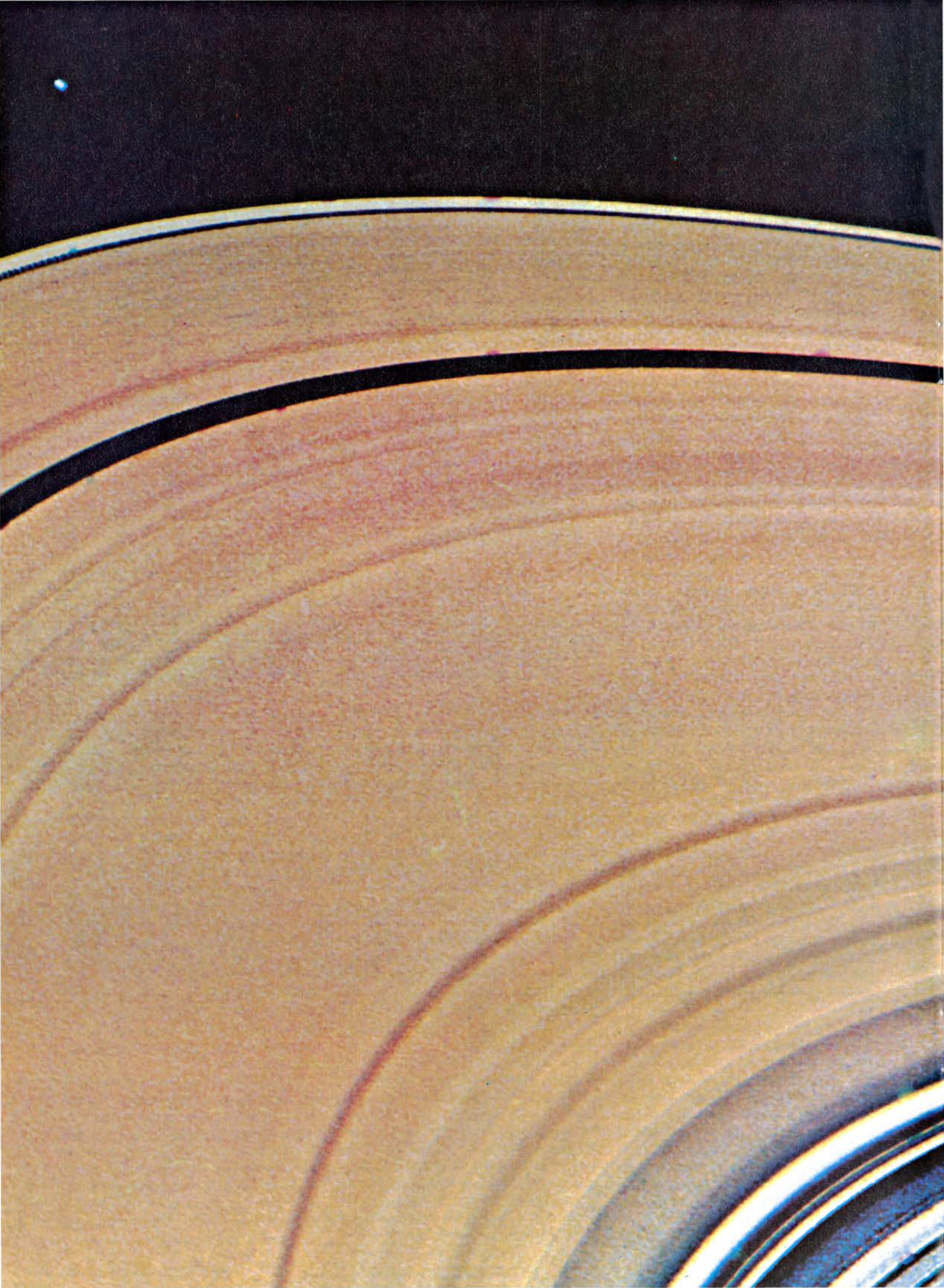
los sacrificios de animales, los había también de frutos y de bebidas o líquidos, cuyos recipientes habían de ser inutilizados, para que no se hiciese de ellos ningún uso profano una vez consagrados. Por último, los objetos de valor y de adorno serían ofrendas o exvotos.

El hecho de que dos de estos "lugares de holocausto" –Langacker y Eisenbichl– se encuentren a sólo 500 metros de distancia revela la gran proximidad a que podían hallarse dos santuarios. Sólo era menester la "señal anómala" de la naturaleza: una piedra de forma extraña, como la Findlingsstein en el Eggli, cuyo nombre actual ("roca del expósito") es aún muy expresivo, un manantial de agua salada (Dürrnberg), la cima de una montaña muy llamativa (Schlern), etcétera.

Las localidades estudiadas por W. Krämer, a quien se debe el artículo que alertó a los investigadores hacia este tipo de yacimientos [W. Krämer, "Prähistorische Brandopferplätze", en *Helvetia Antiqua, Festschrift Emil Vogt*, Zürich, 1966, 111-122] se encuentran en una zona que se extiende desde el alto Danubio hasta la vertiente septentrional de los Alpes y el Schlern, junto a Bozen, en sentido norte sur, y desde Salzburgo hasta el Bodensee en sentido este-oeste. Otros trabajos posteriores han acortado la distancia que separa los "lugares de holocausto" centroeuropeos de los "altares de ceniza" griegos, señalando otros análogos al borde del Adriático y en otros puntos más orientales [M. Menke "Brandopferplatz auf der Kastelliernekropole von Pula, Istrien", *Germania* 48 (1970) 115-123; M. Hell, "Die urzeitliche Opferstelle in Morzg bei Salzburg", *Archaeologia Austriaca* 55 (1974) 100 ss.].

Creemos con todo esto haber señalado una de las vías que la investigación de Cancho Roano debe seguir. No tardarán en detectarse otros lugares afines, o en que se interpreten como tales algunos yacimientos conocidos y explorados antes en el valle del Guadalquivir y del Guadiana.

En último término, lo que estos grandes crematorios acreditan, por encima de todo, es la confianza de los pueblos indoeuropeos en el poder del fuego como vehículo de la plegaria, como elemento trasmisor entre la tierra y el cielo. Esa es la pavorosa lección de Cancho Roano, lo que deja en el ánimo del visitante una impresión imborrable. (Antonio Blanco Freijeiro.)



Anillos en el sistema solar

Tres de los planetas gigantes poseen anillos. Los que ciñen a Saturno constan, a su vez, de miríadas de otros más finos. ¿Cómo conservan su forma anular? Gracias a una complicada interacción de las fuerzas que los esculpen

James B. Pollack y Jeffrey N. Cuzzi

El acercamiento del vehículo espacial *Voyager 2* hasta 100.000 kilómetros de Saturno, el 25 de agosto del año pasado, culminó un período de exploración planetaria, en el que los anillos de Saturno nos han sorprendido quizá tanto como al primero que los vio, hace de ello casi cuatro siglos. Hoy día, los anillos de Saturno se nos muestran con gran riqueza de detalles: bandas, radios y trenzados. Hay aspectos por explicar todavía. Por otra parte, resulta que Saturno no es el único planeta dotado de anillos. Uno tiene Júpiter y nueve anillos discretos, al menos, Urano. Queda por determinar si los presenta Neptuno y si, por tanto, los anillos son ubicuos entre los planetas gaseosos gigantes del sistema solar exterior. Aquí trataremos de la estructura y composición de los anillos de Júpiter, Saturno y Urano, con especial atención a los del segundo citado. Nos ocuparemos luego del estado actual del conocimiento sobre los procesos que dan forma a los anillos. Para abordar, por último, las distintas explicaciones posibles de cómo llegaron a formarse.

Unos apéndices extraños

Los anillos de Saturno se observaron por primera vez en julio de 1610. Le cupo tal mérito a Galileo Galilei. En parte porque las imágenes que daba su invento, el telescopio astronómico, eran de mala calidad, y en parte porque hacía sólo unos meses que había descubierto los cuatro mayores satélites de Júpiter, pensó inicialmente que las estructuras borrosas, parecidas a orejas, que había visto eran dos satélites próximos a Saturno. Pronto cambió de opinión. Aquellos "extraños apéndices" no variaban su posición respecto a Saturno de una noche a la siguiente. Y, además, desaparecieron en 1612. ¿Qué había ocurrido? Lisa y llanamente: los anillos habían quedado orientados con su plano según la visual desde la Tierra en 1612 y con ello se habían hecho muy débiles.

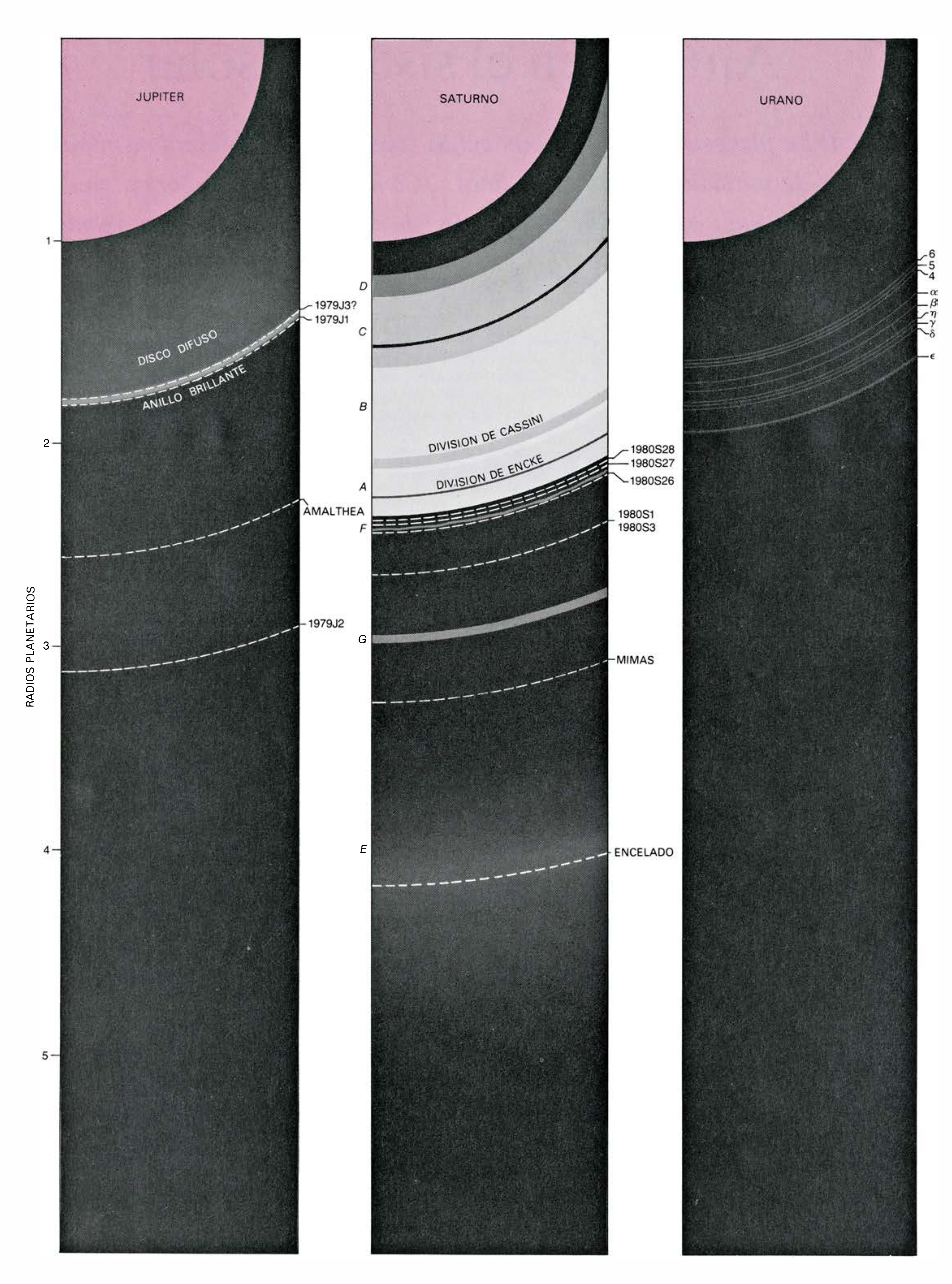
La geometría de los apéndices dejó perplejos a los astrónomos. Hasta el punto de llegarse a proponer que se trataba de asas unidas a Saturno o que constaban de varios satélites en órbita solamente alrededor de la parte posterior de Saturno, por lo que nunca arrojaban sombra sobre el planeta. Finalmente, en 1655, Christiaan Huygens sugirió que los apéndices eran el signo visible de un disco de materia delgado y plano, separado del planeta y dispuesto en el plano ecuatorial de éste. Dependiendo de cuáles fueran las posiciones de Saturno y de la Tierra en sus respectivas órbitas alrededor del Sol, la inclinación del disco respecto a la Tierra variaría; de ahí que su apariencia variase también desde la de una delgada línea hasta la de una ancha elipse.

Durante los dos siglos siguientes se supuso que el disco era una capa continua de materia. La primera objeción contra la hipótesis no tardaría, sin embargo, en plantearse. En 1675, Jean Dominique Cassini halló una oscura banda –la división que lleva su nombre– que separaba el disco en dos anillos concéntricos. Asimismo, a finales del siglo XVIII, Pierre Simon de Laplace mostró que bastarían las fuerzas combinadas de la gravedad de Saturno y la rotación del disco para desgarrar una capa única de materia. En principio, cualquier partícula del disco mantiene su distancia radial desde Saturno porque hay dos fuerzas que se equilibran. La gravedad tira de la partícula hacia dentro; la fuerza centrífuga la empuja hacia fuera. La fuerza centrífuga procede de la velocidad de rotación; de aquí que el disco haya de estar girando. Ahora bien, en el caso de un disco en rotación rígida, las fuerzas se equilibran solamente para una cierta distancia radial. Por ello, Laplace propuso la hipótesis de que los anillos de Saturno estaban formados por muchos anillos delgados, cada uno de ellos suficientemente delgado para soportar el ligero desequilibrio de fuerzas que aparecería a lo largo de su anchura radial.

El último paso hacia la visión moderna de los anillos se dio en 1857, cuando James Clerk Maxwell ganó el premio Adams de la Universidad de Cambridge por su demostración matemática de que los anillos delgados estaban formados en realidad por numerosas masas pequeñas que mantenían órbitas independientes. La comprobación experimental de esta hipótesis llegó desde distintos flancos. Así, en 1895, los astrónomos americanos James E. Keeler y William W. Campbell dedujeron la velocidad de las partículas en los anillos a partir de sus corrimientos Doppler, o alteración de la longitud de onda de las líneas espectrales de la luz del Sol que las partículas reflejan hacia la Tierra. Encontraron que los anillos giraban alrededor de Saturno a una velocidad distinta de la de la atmósfera del planeta. Además, las partes internas de los anillos giraban a mayor velocidad que las externas, según prescribían las leyes de la física para partículas en órbitas independientes.

Los anillos de Urano se descubrieron por azar. Había varios grupos de astró-

ANILLO "A" DE SATURNO, fotografiado el 23 de agosto por el vehículo espacial *Voyager 2*. En la imagen resultante, presentada en falso color, el anillo *A* es amarillento; el intervalo más ancho del anillo corresponde a la división de Encke. La parte del anillo exterior a la división de Encke muestra un débil grupo de bandas. Las bandas están más apretadas hacia la órbita de 1980S27, satélite que se descubrió en las imágenes tomadas por el *Voyager 1* y visible como una mancha blanca en la parte superior izquierda. Se cree que las bandas se producen por resonancias en el anillo debidas a los efectos gravitatorios del satélite. Además, 1980S27 y un segundo satélite, 1980S26, son los que dan forma al anillo *F* de Saturno, demasiado débil para verse aquí. Los pequeños anillos blanquecinos y azulados (*abajo, a la derecha*) ocupan la división de Cassini. Sus diferentes colores revelan una distinta composición de sus partículas.


JUPITER
SATURNO
URANO
RADIOS PLANETARIOS
1
2
3
4
5
1979J3?
1979J1
DISCO DIFUSO
ANILLO BRILLANTE
AMALTHEA
1979J2
D
C
B
A
F
G
E
DIVISION DE CASSINI
DIVISION DE ENCKE
1980S28
1980S27
1980S26
1980S1
1980S3
MIMAS
ENCELADO
6
5
4
α
β
η
γ
δ
ε

nomos con todo preparado para observar el paso de la estrella SAO 158687 por detrás de Urano el 10 de marzo de 1977. El objetivo: estudiar la estructura de la atmósfera de Urano. Llegado el momento, algunas de las mejores observaciones correspondieron a las realizadas por James L. Elliot y sus colaboradores en el Observatorio Volante Kuiper, que era un avión equipado con un telescopio de 91 centímetros. Su grupo (y varios otros) hallaron que el brillo de la estrella disminuía no sólo cuando pasaba por detrás de Urano, sino también en cierto número de lugares situados cerca del planeta, muy por encima de su atmósfera. Las atenuaciones de corta duración definían una serie de distancias a un lado de Urano que eran aproximadamente simétricas con la serie al otro lado del planeta. Se supuso que la simetría provenía de la presencia de anillos relativamente opacos, muy estrechos y casi circulares. Observaciones posteriores han revelado hasta ahora la presencia de nueve anillos, todos ellos situados dentro de la distancia de un radio planetario contado desde la cima de la atmósfera de Urano.

En el descubrimiento de los anillos de Júpiter intervinieron vehículos espaciales. La primera indicación llegó cuando el *Pioneer 10* pasó cerca de Júpiter en 1974. Júpiter tiene un campo magnético que atrapa las partículas cargadas en ciertas regiones del espacio que ciñe al planeta. Esas regiones vienen a ser el equivalente joviano de los cinturones de Van Allen en la Tierra. Lo que el *Pioneer 10* detectó era un decrecimiento en la cuenta de partículas de alta energía en los cinturones jovianos entre 50.000 y 55.000 kilómetros por encima de la atmósfera del planeta. Mario H. Acuña y Norman F. Ness, del Centro de Vuelos Espaciales Goddard de la Administración Nacional de Aeronáutica y del Espacio (NASA), sugirieron que el decrecimiento se podía deber a absorción parcial de partículas por un satélite próximo o un sistema de anillos. Lo último resultó ser el caso. En 1979 se detectó un débil anillo joviano cuando fue fotografiado por los sistemas de elaboración de imágenes de los *Voyager 1* y *Voyager 2*. Los vehículos espaciales fueron desviados hacia Saturno por el campo gravitatorio de Júpiter. Llegaron a la proximidad de Saturno en noviembre de 1980 y agosto de 1981, respectivamente.

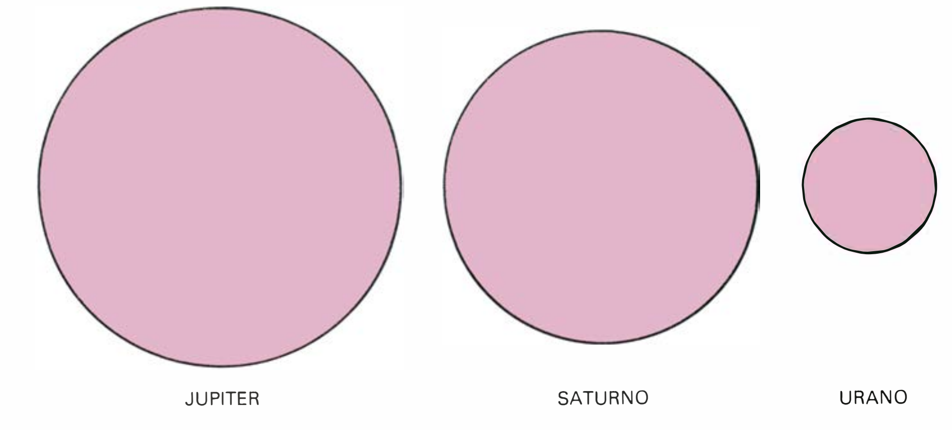
Características de los anillos planetarios

Los anillos de Saturno, Urano y Júpiter comparten cierto número de propiedades. En primer lugar, están compuestos de miríadas de partículas en órbitas independientes. En segundo lugar, se encuentran mucho más cerca del planeta padre que cualquiera de sus satélites principales; de hecho, el grueso de cada sistema de anillos se encuentra a una distancia de la superficie del planeta inferior a un radio planetario. En tercer lugar, los anillos están situados en el plano ecuatorial del planeta; de hecho, casi toda la materia que constituye los anillos se halla confinada en una delgada región en ese plano. En cuarto lugar, el sistema de anillos de Júpiter y Saturno tiene cierto número de pequeños satélites cerca o dentro de los anillos. Se sospecha la existencia de pequeños satélites semejantes, asociados con los anillos de Urano.

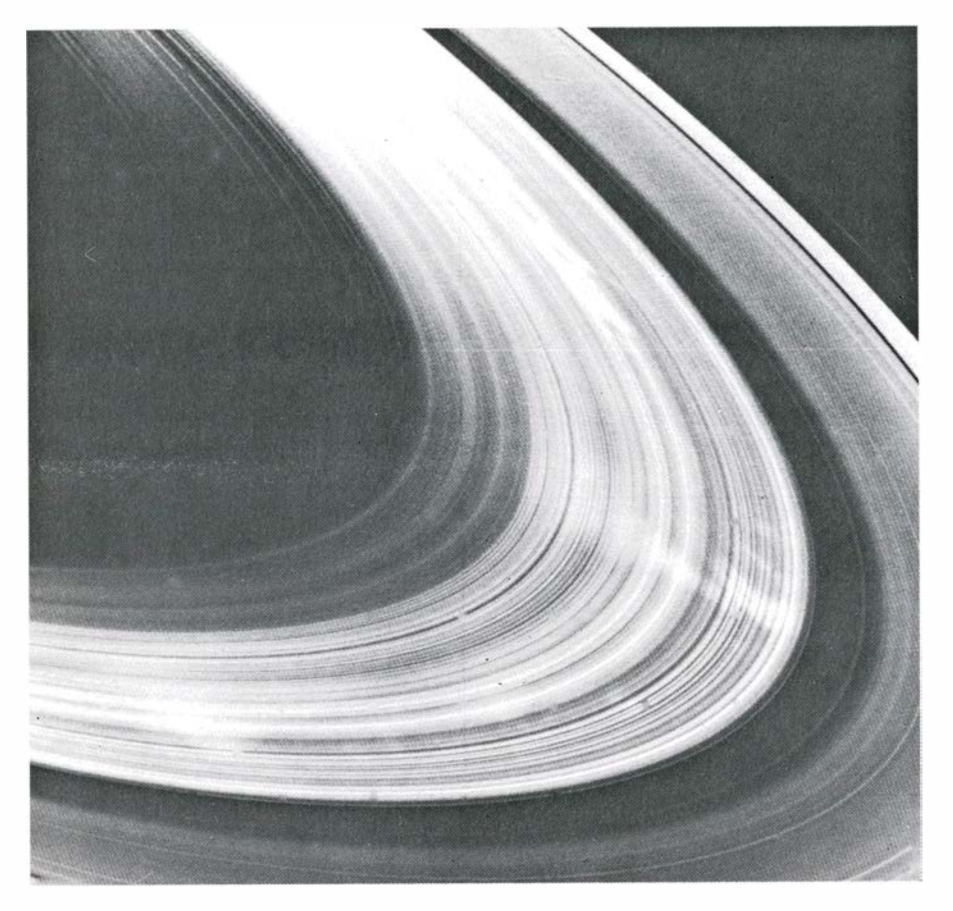
Pero cada sistema de anillos muestra sus propias peculiaridades. El de Saturno tiene siete porciones importantes. Algunas de ellas están separadas de las porciones vecinas por espacios anulares más o menos vacíos; los bordes de los otros se caracterizan por cambios en la densidad de la distribución de partículas de anillos. Cada porción o sección se designa por una letra, que refleja no su distancia a Saturno, sino el orden en que se descubrieron o se postularon las secciones.

El cuerpo principal del sistema de anillos de Saturno incluye, pues, los brillantes anillos *A* y *B*, de escasa opacidad. Media entre uno y otro una franja de 5000 kilómetros: la división de Cassini, región relativamente transparente, aunque no vacía en absoluto. El cuerpo principal del sistema de Saturno comprende también el anillo *C*, más débil y menos opaco, que queda dentro del borde interior del anillo *B*. Tiene un grado de opacidad comparable al de la división de Cassini. El todavía más débil anillo *D* queda dentro del anillo *C*. Hay finalmente tres anillos muy pálidos, *E, F* y *G*, que quedan fuera del anillo *A*. Tomados en su conjunto, los principales anillos de Saturno (los *A, B* y *C*) miden unos 275.000 kilómetros de anchura anular, lo que representa tres cuartas partes de la distancia que separa la Tierra de la Luna. En comparación, el espesor de los anillos de Saturno resulta despreciable. El límite superior de su extensión vertical se ha estimado en alrededor de un kilómetro. En relación con su anchura, los anillos son miles de veces más delgados que una hoja de afeitar.

La capacidad de los anillos para reflejar o absorber luz de diferentes longitudes de onda permite deducir información sobre la composición de las partículas de los anillos de Saturno. Por ejemplo, los anillos *A, B* y *C* son malos reflectores de la luz del Sol para ciertas longitudes de onda del infrarrojo próximo. Por tratarse de una propiedad característica del hielo, cabe presumir que el hielo es un constituyente importante de las partículas que forman esos anillos. Pero es un hielo de color blanco, lo que significa que es más o menos igualmente reflector para todas las longitudes de onda en el visible. Por el contrario, las partículas de los anillos



SISTEMAS DE ANILLOS en torno a Júpiter, Saturno y Urano, esquematizados a escalas en las que se toma como unidad el radio de cada planeta. Júpiter (*izquierda*) tiene un anillo "brillante" que, en realidad, es muy débil y casi transparente. Hacia el interior del anillo se extiende un disco de partículas todavía más débil, que quizás alcanza hasta la atmósfera del planeta. Un halo de partículas (no representado en esta ilustración) confiere al sistema un espesor vertical de unos 20.000 kilómetros. El sistema se descubrió gracias a los vehículos espaciales Pioneer y Voyager. Saturno (*centro*) está ceñido por siete anillos. Los anillos *A* y *B* están separados por la división de Cassini; el anillo *A* incluye la división de Encke. Las letras se asignaron a los anillos en el orden de su descubrimiento. Sólo los anillos principales (*A, B* y *C*) se ven fácilmente mediante telescopios emplazados en la Tierra; el resto (excepto el anillo *E*) se descubrió gracias a vehículos espaciales. Urano (*derecha*) tiene no menos de nueve anillos. Aquí se ha exagerado su anchura para distinguirlos. Se detectaron desde la Tierra y se designan por números o letras griegas. Los satélites del interior del sistema de anillos se rotulan a la derecha de cada cuadro; sus órbitas se representan mediante líneas de trazos. En esta página se muestran los tamaños relativos de Júpiter, Saturno y Urano.





CUÑAS RADIALES Y BANDAS de los anillos de Saturno, fotografiadas por primera vez por el *Voyager 1*. Las cuñas radiales aparecen esporádicamente en el anillo *B*. Cada una de ellas es brillante en las imágenes de la cara de los anillos iluminada por el Sol, tomadas cuando el vehículo espacial y el Sol están en lados opuestos de Saturno, de modo que la luz del Sol se refleja hacia delante desde las partículas del anillo hasta el vehículo espacial (*fotografía de arriba*); cada cuña aparece oscura en las imágenes tomadas cuando el vehículo espacial y el Sol se encuentran en el mismo lado de Saturno, de modo que la luz solar se refleja hacia atrás (*fotografía de abajo*). En ambas imágenes es evidente la estructura de los anillos formando bandas y anillos delgados. La banda oscura y ancha recibe el nombre de división de Cassini.

A, *B* y *C* son menos reflectoras en luz azul que en luz roja. Quizás hay alguna substancia adicional presente en pequeñas cantidades; polvo tal vez, que portara óxido de hierro como fuente del color rojizo. También se ha propuesto la hipótesis de que ciertos compuestos generados por la radiación ultravioleta del Sol fueran los responsables del color rojizo. Algunos compuestos incoloros que contienen azufre se convierten en polisulfuros bajo la acción de la radiación ultravioleta; los polisulfuros absorben selectivamente en el azul. Los datos suministrados por los Voyager revelaron algo sorprendente: las partículas de los anillos *A* y *B* tienen un color semejante a las del anillo *C* y de la división de Cassini, si bien son más brillantes y más rojizas.

Las observaciones con radar permiten avanzar más en el razonamiento. En 1973, Richard M. Goldstein y Gregory Morris, del Laboratorio de Propulsión a Chorro del Instituto de Tecnología de California, exploraron los anillos de Saturno con ondas de radar cuya reflexión detectaron con la antena de 64 metros de la Red de Espacio Profundo en Goldstone, California. La alta reflectividad de los anillos *A* y *B* implicaba que la mayoría de las partículas de esos dos anillos eran al menos de un tamaño comparable a la longitud de onda del radar empleado por los investigadores; medía éste varios centímetros. Si hubieran sido menores que las longitudes de onda del radar, habrían resultado transparentes a las ondas de éste. Si hubieran sido mucho mayores, se habría apreciado su emisión de radiación térmica a esas longitudes de onda. El bajo nivel de tal radiación limita su tamaño a no más de algunos metros.

Los datos de los vehículos espaciales Voyager han confirmado y ampliado los descubrimientos iniciales. En un tipo de experimento se enviaron radioondas desde el vehículo espacial a la Tierra, a través de los anillos, y se midió la potencia difundida por las partículas de los anillos para varios ángulos de desviación respecto al trayecto inicial de las ondas. A medida que el tamaño de las partículas aumenta con relación a la longitud de onda, la curva polar de dispersión queda más estrechamente confinada a pequeños ángulos de desviación respecto a la dirección de propagación. De acuerdo con el análisis de los datos de los Voyager realizado por G. Leonard Tyler y Ahmed Essam A. Marouf, de la Universidad Stanford, las mayores entre las partículas que abundan en los anillos *A*, *B* y *C* tienen un diámetro de unos 10 metros. Menudean, todavía más, partículas de

menor tamaño, hasta 10 centímetros, y existen variaciones regionales en la distribución de tamaños dentro de los anillos.

Así como la difusión de las ondas de radar por las partículas en los anillos hace posible detectar partículas del orden del tamaño de la longitud de onda del radar, la difusión de la luz solar permite detectar partículas del tamaño de una longitud de onda de la luz visible. El intenso incremento de brillo de un segmento del anillo, cuando se contempla bajo un ángulo para el que la difusión hacia delante es pequeña, implica que, en ese segmento, abundan las partículas de un micrometro de magnitud. Observación que sólo puede acometerse cuando Saturno queda entre el Sol y el astrofísico. Esta condición no se puede cumplir para observaciones verificadas desde la Tierra, pero sí a bordo de un vehículo espacial. Así, los estudios de los datos de los Voyager señalan que las partículas de tamaños del orden de un micrometro constituyen una proporción grande de las partículas en el anillo *F*, una proporción apreciable en muchas partes del anillo *B* y una proporción menor en la parte externa del anillo *A*. Por otra parte, el anillo *C* y la división de Cassini no presentan rastros de tales partículas pequeñas.

Estructura de los anillos de Saturno

Antes de que los Voyager pasaran por la proximidad de Saturno se había reconocido ya cierta configuración estructural en los anillos del planeta. Hemos hablado de la división de Cassini y se conocía también la división de Encke, una banda más estrecha en la parte externa del anillo *A*. Las fotografías de los anillos con alta resolución, tomadas por los vehículos espaciales Voyager, aportaron más de una sorpresa. Aparecieron estrechas regiones anulares de diferente brillo y opacidad, como los surcos del disco de un gramófono. Se hallaron, además, desviaciones respecto a la forma circular. Incluían éstas unas prolongaciones orientadas radialmente, en forma de cuña, en el anillo *B* y nudos, trenzados y torcimientos en el anillo *F*.

Donde se advierte una mayor organización estructural es en el anillo *B*, que posee también la mayor densidad de partículas que se encuentra en los anillos. Hay variaciones de opacidad en el anillo *B* a lo largo de distancias radiales de no más de 10 a 50 kilómetros; lo que no deja de llamar la atención si consideramos que posee una anchura total de 25.000 kilómetros. La parte central y más opaca de este anillo *B* es aquella donde aparecen las cuñas radiales, cada una de las cuales puede verse a lo largo de una fracción importante de las 10 horas que una partícula del anillo *B* invierte en realizar una revolución orbital. Mientras tanto, nuevas cuñas radiales están surgiendo esporádicamente en otras zonas del anillo. Comparadas con su entorno, las cuñas radiales aparecen brillantes en luz difundida hacia delante y oscuras en luz difundida hacia atrás. De ahí que las partículas de tamaños del orden de un micrometro abunden en las cuñas radiales.

Cada parte de una cuña radial gira alrededor de Saturno a la misma velocidad que lo hacen las partículas del anillo a su distancia radial. Las porciones interiores se mueven más aprisa; así, una cuña radial se va inclinando con el tiempo y llega a desaparecer. El extremo más estrecho (el "pincho") de cada cuña radial parece coincidir aproximadamente con la distancia desde Saturno a la cual el período de una partícula en órbita iguala al período de rotación de Saturno. El campo magnético de Saturno está encerrado dentro del planeta; gira, por tanto, con él. De aquí que las fuerzas electromagnéticas sean parcialmente responsables de que existan cuñas radiales. A este respecto puede hacerse notar que en el experimento de Radioastronomía Planetaria a bordo del *Voyager 1* se observaron brotes de estática de banda ancha. Los brotes parecen haberse originado en fuentes del anillo *B*, cerca de regiones donde la actividad de las cuñas radiales era intensa.

Si lo comparamos con el anillo *B*, el *A* no muestra rasgos que se destaquen y brilla uniformemente, si bien presenta varias características distintivas, estrechas y discretas, en sus regiones más externas. Por una parte, la banda de la división de Encke, oscura y relativamente vacía, con una anchura de unos 350 kilómetros, contiene dos o tres anillos irregulares, pequeños y estrechos, formados por partículas pequeñas. Las imágenes tomadas por el *Voyager 2* muestran que esos anillos de escaso tamaño poseen arrugas o pliegues que recuerdan los detectados por el *Voyager 1* en el anillo *F*. La división de Encke está confinada por cuatro bandas poco corrientes, con estructura irregular y ondulada. A mayor distancia de Saturno, el borde exterior del anillo *A* presenta muchas bandas estrechas, de anchura próxima a los 20 kilómetros, en las que abundan partículas pequeñas. El espaciado entre las bandas decrece hacia el exterior. El borde externo del anillo *A* está muy definido. También aquí proliferan partículas pequeñas.

El anillo *C* y la división de Cassini presentan muchas semejanzas estructurales, amén de su relativa falta de color rojo, grado parecido de transparencia y escasez de partículas pequeñas. Ambos manifiestan bandas discretas, regularmente espaciadas y de brillo uniforme. Uno y otra contienen también huecos estrechos, de bordes bien definidos y completamente vacíos con una anchura radial de 50 a 350 kilómetros. Algunos de los huecos poseen, además, pequeños anillos más delgados aún, de bordes igualmente bien definidos y muy opacos. Varios de estos últimos anillos son excéntricos, vale decir, no circulares, y de anchura no uniforme. Las imágenes del *Voyager 2* muestran que los anillos opacos difieren frecuentemente en color del material de los anillos circundantes. Se parecen más a la materia de los anillos *A* y *B*.

Anillos de Urano y Júpiter

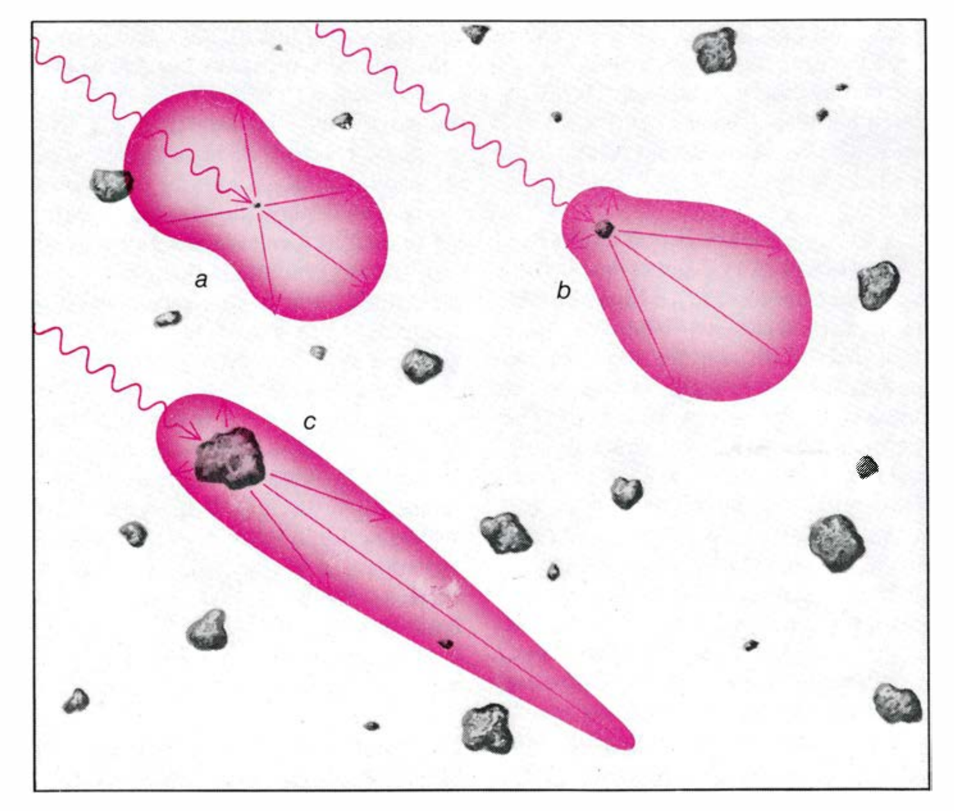
Los anillos de Urano complementan, en muchos aspectos, los de Saturno. Los principales anillos de este planeta (los *A*, *B* y *C*) son anchos y están intercalados con estrechos huecos; los de Urano, por el contrario, son estrechos (miden desde algunos kilómetros hasta 100 kilómetros de anchura radial) y están separados por amplias regiones vacías. Se ha confirmado la existencia de nueve anillos alrededor de Urano. A cada uno se le ha asignado una letra griega. La técnica de la ocultación estelar, aplicada al estudio de los anillos, aprovecha el pequeño tamaño aparente de la estrella ocultada para proporcionar una alta resolución espacial. Lo que no empece que hayan quedado varios anillos sin resolver; se trata de los que tienen una extensión radial inferior a cinco kilómetros. (La técnica de la ocultación estelar, tal como se ha aplicado a los anillos de Saturno por el *Voyager 2*, revela la estructura en escala muy fina, de cientos de metros.) Los anillos de Urano más estrechos (los designados γ, δ y η) vienen a ser casi circulares y están situados en un plano común; por el contrario, los anillos α, β y 4, algo más anchos, tienden a ser ligeramente elípticos e inclinados respecto al plano común. El anillo ε es el más ancho y elíptico de los nueve. Su anchura radial varía linealmente con la distancia de Urano; empieza con una anchura de unos 20 kilómetros donde el anillo está más próximo al planeta y acaba con una anchura de 100 kilómetros, en el punto de alejamiento máximo. Algunos de los anillos pequeños y estrechos de Saturno exhiben, como se sabe ahora, una variabilidad semejante.

La atenuación del brillo de las estrellas al pasar por detrás de la posición de los anillos de Urano indica que éstos revelan una opacidad comparable con las de las partes más opacas (y por ende más brillantes) de los anillos de Saturno. El anillo ε es particularmente opaco. Cosa curiosa, presenta su mayor opacidad en los bordes. Las fuerzas gravitatorias que provienen de la forma aplastada de Urano hacen que la orientación de los anillos elípticos, ε incluido, cambien de un modo incesante. Un ciclo durante el cual la forma elíptica del anillo gira una vez alrededor del planeta tarda en completarse unas 6300 horas. En ese tiempo, una partícula en los anillos describe unas 750 órbitas alrededor de Urano. De aquí que haya dos cosas que afectan al anillo ε. En primer lugar, cambia su orientación. En segundo lugar, las partículas del anillo sitas a diferentes distancias radiales de Urano describen órbitas en torno al planeta a velocidades algo diferentes, de acuerdo con la física de los movimientos orbitales. Lo que no obsta para que el anillo ε mantenga su integridad. Es decir, persisten su excentricidad, su anchura variable y la concentración de partículas hacia sus bordes.

¿De qué están compuestos los anillos de Urano? Aunque no es fácil detectarlos desde la Tierra por causa de su carácter angosto, observaciones recientes ponen de manifiesto que las partículas del anillo son muy oscuras. Ya no es aquí componente dominante el hielo. Con toda probabilidad, las partículas del anillo en órbita alrededor de Urano están constituidas por silicatos ricos en compuestos que absorben la luz solar. Ciertos óxidos de hierro y compuestos carbonados complejos satisfacen las características requeridas. Por desgracia, nada sabemos acerca del tamaño de las partículas.

Los anillos de Júpiter constan de tres partes principales: un anillo brillante, un disco difuso y un halo. El anillo brillante tiene una anchura de unos 6000 kilómetros. Su borde exterior, bastante definido, se encuentra a unos 58.000 kilómetros, o 0,8 radios jovianos, por encima de la superficie de Júpiter. En la parte externa del anillo hay una estrecha banda de unos 600 kilómetros de anchura, que es un 10 por ciento más brillante que el resto. Empero, la opacidad del anillo "brillante" es tan baja que sólo el 0,001 por ciento de la luz solar que pasa a través de él es interceptada por sus partículas. El disco difuso es varias veces más débil. Se extiende hacia el interior desde el margen interno del anillo brillante. Puede llegar incluso a alcanzar la atmósfera de Júpiter. Vistos de canto, el anillo brillante y el disco parecen estar confinados principalmente a un espesor no superior a 30 kilómetros. Cabe destacar, sin embargo, que el halo posee una extensión vertical de unos 20.000 kilómetros. El halo presenta su máximo espesor sobre el disco difuso. Su límite exterior se prolonga un poco más allá del borde exterior del anillo brillante.

Basándose en la manera como el anillo brillante aumenta su resplandor a pequeños ángulos de difusión, se ha descubierto que contiene partículas cuyo tamaño característico es de varios micrometros. Tales partículas son muy poco eficaces para absorber protones y electrones de alta energía; no podrán ser, pues, las responsables del decrecimiento en el flujo de partículas de alta energía que ha detectado, en el anillo brillante, el *Pioneer 10*. Debe haber también partículas de un centímetro de diámetro, por lo menos. Las partículas del anillo son rojizas, lo que las hace del mismo color de muchos asteroides y satélites del sistema solar exterior.



LA DIFUSION DE LA LUZ o de alguna otra forma de radiación electromagnética por las partículas de un anillo permite deducir el tamaño de las partículas que abundan en el anillo. En particular, si una partícula tiene una dimensión menor que aproximadamente una décima parte de la longitud de onda de la radiación incidente, difunde la luz casi por igual en todas las direcciones (*a*). Si la partícula es algo menor que la longitud de onda, tiende a difundir la radiación hacia delante (*b*). Si la partícula es mayor que la longitud de onda, difunde hacia delante casi toda la radiación (*c*). La longitud de las flechas representa la cantidad relativa de energía difundida a varios ángulos. La observación de que las partículas de las cuñas radiales del anillo *B* de Saturno difunden la luz predominantemente hacia delante permite deducir que las cuñas son concentraciones locales y transitorias de partículas del anillo, de un micrometro de tamaño.

Choques en los anillos

La arquitectura de un sistema de anillos resulta de la interacción de cierto número de fuerzas. Se encuentran, entre éstas, las fuerzas gravitatorias debidas a los satélites exteriores a los anillos y a los pequeños satélites inmersos en ellos, las fuerzas electromagnéticas debidas al campo magnético en rotación del planeta e incluso las débiles fuerzas ejercidas por el medio gaseoso enrarecido donde giran los anillos.

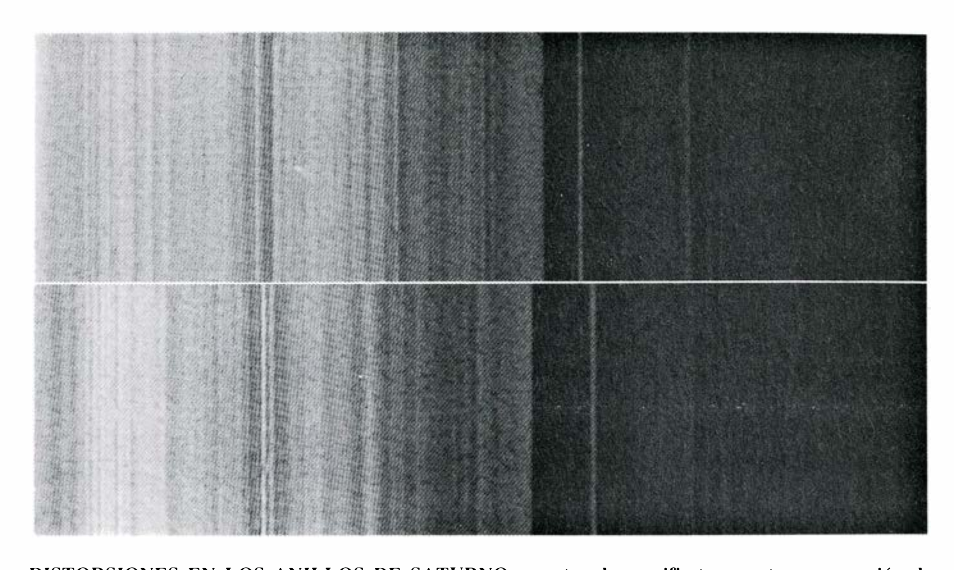
Todas las partículas de un sistema de anillos comparten un movimiento orbital común en torno al planeta: viajan en la dirección de la rotación de éste. Los movimientos vertical y radial superpuestos al movimiento orbital de cada una de tales partículas no están sujetos a esa ligadura. De aquí que las partículas vecinas se muevan arbitrariamente en esas direcciones con respecto a las otras partículas; y así los choques resultan inevitables. Cuando las velocidades relativas aleatorias son grandes, como podría ocurrir en el caso de que los anillos fuesen una densa nube de partículas, se producen choques violentos, y aun cuando sean raros, se invierte una gran cantidad de la energía del movimiento relativo en calentar las partículas y deformar su estructura. La consiguiente pérdida de energía significa que

las velocidades aleatorias decrecen rápidamente. El decrecimiento de la componente vertical de las velocidades conduce a un aplastamiento del sistema de anillos. Al propio tiempo, el decrecimiento en la componente radial conduce a órbitas más circulares. En resumen, un anillo grueso se convierte en un disco delgado y aproximadamente circular en una fase muy temprana de su historia.

No importa que las partículas del anillo hayan perdido casi todo su movimiento aleatorio: los choques continúan. La razón de ello es la siguiente: la fuerza gravitatoria ejercida por el planeta sobre las partículas de un anillo se debilita conforme aumenta la distancia desde el planeta, de modo que las partículas del anillo situadas a mayor distancia tardan más en completar su circunvalación en torno al planeta. Así, una de esas partículas cuya órbita esté ligeramente por dentro de la de una segunda partícula del mismo anillo llegará, con el tiempo, a alcanzarla y las dos chocarán si la separación radial que media entre ellas es menor que el diámetro de una partícula.

Es probable que el choque tenga lugar a una velocidad relativa de menos de un centímetro por segundo. Pero puede convertir una fracción del movimiento orbital circular de la partícula en movimiento vertical aleatorio. Los subsiguientes choques impedirán que las velocidades verticales de las partículas resulten muy violentas. Se alcanzará, por tanto, un régimen estacionario que determinará el espesor del anillo. Si las partículas tienen un amplio espectro de tamaños, las más pequeñas ganarán velocidad vertical, sobre todo, al ser desviadas gravitatoriamente en casichoques con otras partículas mayores. Mas perderán velocidad vertical por choque con otras partículas pequeñas, principalmente. Bajo estas condiciones, las pequeñas alcanzarán una extensión vertical de varias veces el tamaño de las partículas abundantes más grandes. En el caso de los anillos *A* y *B* de Saturno, sería de esperar que las pequeñas partículas ocupasen una extensión vertical de 10 a 100 metros. Mediciones realizadas en el curso de experimentos desde los Voyager indican que los anillos principales de Saturno tienen un espesor que no supera algunos cientos de metros. Este intervalo hace compatibles los espesores observados y predichos.

El choque de partículas vecinas convertirá también parte de su movimiento circular en movimiento radial. Los anillos se extenderán, pues, radialmente. Un anillo aislado y no sometido a acciones exteriores se expandirá hasta que las partículas que los componen se encuentren lo suficientemente alejadas entre sí para que los choques terminen por cesar. El anillo brillante de Júpiter puede haber alcanzado ese estado final; su anchura y opacidad, muy baja, pueden reflejar la expansión, inducida por los choques, de sus partículas mayores a lo largo de la vida del sistema solar.



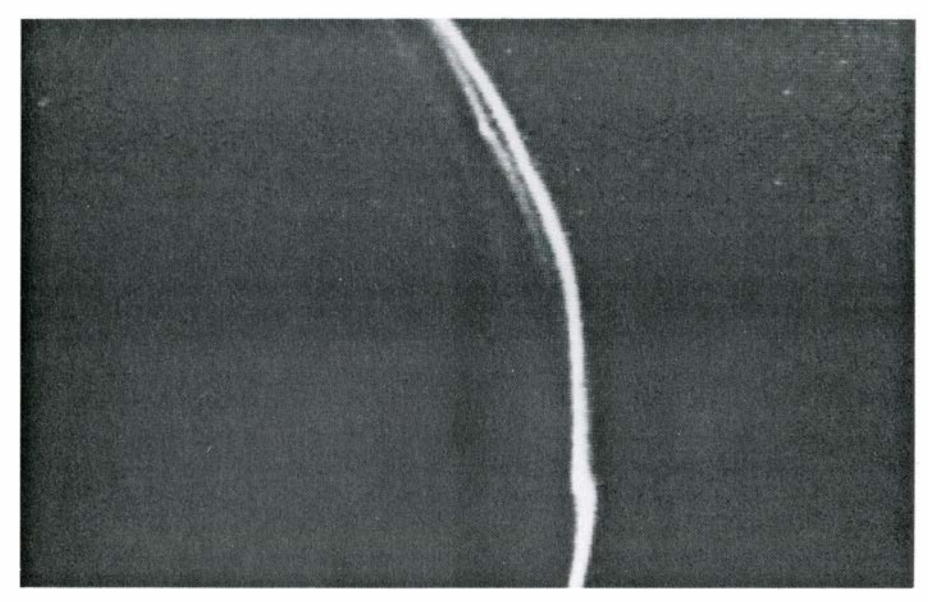
DISTORSIONES EN LOS ANILLOS DE SATURNO, puestas de manifiesto en esta comparación de imágenes tomadas en lados opuestos de los anillos por el *Voyager 2*. En cada imagen, el anillo *B* es la región brillante de la izquierda; la división de Cassini, la región oscura a la derecha. Puede apreciarse que el borde exterior del anillo *B* no es circular, y se confirma la presencia de un delgado anillo brillante y excéntrico por la parte de fuera del anillo *B*. (El anillo delgado fue descubierto por el *Voyager 1*.) Las partículas de la proximidad del borde exterior del anillo *B* describen órbitas en torno a Saturno dos veces por cada órbita completa del satélite Mimas. Esta resonancia amplifica grandemente el efecto gravitatorio de Mimas y distorsiona, por tanto, el anillo *B*; en conjunto, la resonancia crea protuberancias en lados opuestos del anillo. Por el contrario, el delgado anillo excéntrico tiene un máximo único que puede ser el resultado del campo gravitatorio de un pequeño satélite local que no se ha llegado a detectar todavía.

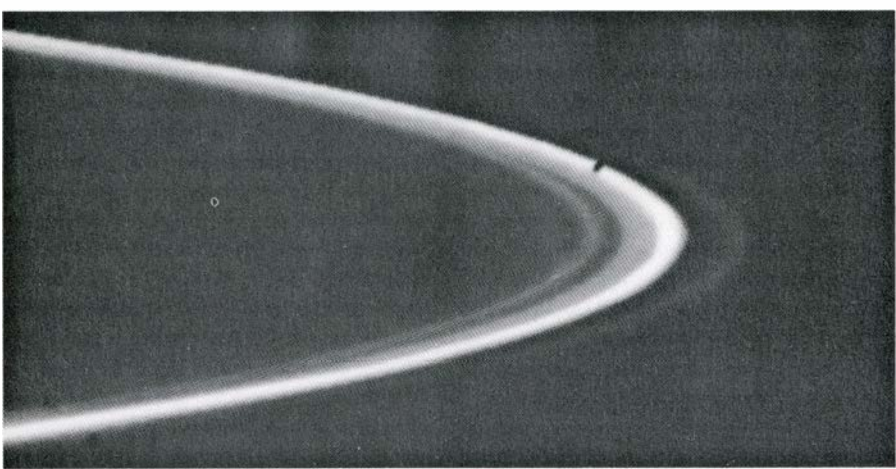
Los anillos de Saturno y Urano muestran, sin embargo, bordes muy marcados que limitan regiones densamente pobladas de partículas. Debe haber, por tanto, otros procesos que contrarresten la rápida expansión inducida por los choques. En esos procesos, los pequeños satélites inmersos en los anillos, o adyacentes a ellos, pueden desempeñar un papel importante. El campo gravitatorio de los satélites mayores y más distantes quizá sirva para encerrar algunos de los pequeños satélites locales en órbitas fijas e impedir de esa manera que se despedacen por causa de la interacción gravitatoria con las partículas del anillo situadas a su alrededor.

El movimiento de los cuerpos en órbita alrededor de un planeta de masa mucho mayor está dominado, fundamentalmente, por el campo gravitatorio de éste. Pero en algunos casos la atracción gravitatoria entre dos de esos cuerpos relativamente pequeños en órbita puede intensificarse y llegar a afectar de forma notable sus movimientos. Amplificación de la fuerza que recibe el nombre de resonancia. Considérese que, si el período orbital de un satélite es un múltiplo exacto o una fracción del de otro satélite, el efecto gravitatorio neto de cada satélite sobre el otro vendrá a ser, en resumidas cuentas, un tirón o un empujón aplicado, repetidamente, en el mismo punto del movimiento cíclico. Así se amplifica el efecto. En algunos casos, la resonancia encierra pares de satélites en órbitas cuyos períodos mantienen una relación fija de dos pequeños números enteros. Tales conmensurabilidades se conocen bien en los sistemas de satélites de Júpiter y Saturno, donde se aplican al menos a cuatro pares de los más importantes satélites y, probablemente también, a varios de los recién descubiertos.

Resonancia y anillos

La resonancia en un disco de partículas da por resultado una situación bastante diferente. Cerca de las distancias radiales del planeta a las que las partículas del disco tendrían un período orbital conmensurado con el de uno de los satélites del planeta, la amplificación del efecto gravitatorio del satélite durante largos períodos de tiempo hace que las órbitas de las partículas dejen de ser circulares. Y aumenta la probabilidad de que las partículas choquen con sus vecinas menos perturbadas. ¿Qué acontece entonces? Se pierden partículas en una banda situada a la distancia radial correspondiente a una resonancia. La banda suele abarcar una anchura natural de unas decenas de kilómetros. En los anillos de Saturno, los





TRENZADOS EN EL ANILLO "F" DE SATURNO, revelados en fotografías hechas por el *Voyager 1* a su paso por la vecindad del planeta en noviembre de 1980; sin embargo, cuando nueve meses más tarde llegó el *Voyager 2*, habían desaparecido. El anillo mismo se halla a unos 80.000 kilómetros (1,3 veces el radio de Saturno) de la superficie de este planeta y a 4000 kilómetros del borde exterior del anillo *A*. En noviembre de 1980 (*arriba*), constaba de tres cintas, cada una de ellas de una anchura de unos 30 kilómetros. Las dos cintas más externas, que mostraban arrugas, torsiones y nudos, parecían trenzadas o al menos cortarse entre sí. Los campos gravitatorios de los satélites designados 1980S26 y 1980S27 pueden explicar tal deformación estructural. En agosto de 1981 (*abajo*), el anillo *F* había cambiado ya de estructura. Predominaba entonces una cinta sin trenzar y aparecían como compañeras otras tres cintas más débiles.

ejemplos de tales bandas incluyen, seguramente, la docena de estrechos huecos en la parte externa del anillo *A*, que al parecer resultan de resonancias producidas por los satélites 1980S1, 1980S3, 1980S26 y 1980S27, recientemente descubiertos. El predominio de pequeñas partículas cerca de los huecos atestigua, verosimilmente, la violencia de los choques inducidos localmente por cada resonancia.

Se ha de mencionar un efecto adicional de las resonancias que están distantes de los pequeños satélites. En 1978, Peter M. Goldreich, del Instituto de Tecnología de California, y Scott D. Tremaine, ahora en el Instituto de Estudios Superiores de Princeton, lanzaron la hipótesis de que en los anillos de Saturno había ondas espirales de fluctuaciones de la densidad de materia en el anillo. Anteriormente se había sugerido que ondas similares eran las responsables de la configuración espiral de los brazos de ciertas galaxias, la Vía Láctea por ejemplo. En imágenes construidas por el *Voyager 1*, parece haber ondas de densidad débilmente visibles en la división de Cassini: su aspecto, estrechamente arrollado, recuerda el del muelle real de un reloj. Se cree que estas ondas se han producido allí por resonancias. Las ondas pueden transportar los efectos de las resonancias hasta grandes distancias; en virtud de lo cual cabe presumir que las ondas desempeñen un papel central en el transporte de materia dentro del disco formado por los anillos de Saturno.

Las resonancias creadas en un disco por un satélite están acortando sus intervalos a medida que nos acercamos a la órbita del satélite. A cierta distancia crítica, la separación radial entre resonancias sucesivas se hace igual a la anchura natural de cada resonancia. Dentro de esta distancia crítica, las resonancias se superponen. El resultado es una zona continua de transporte de materia del anillo que extrae materia de la órbita del satélite. La anchura de la zona y el grado de escasez de partículas en ella dependen de la masa del satélite y de la densidad de materia en el anillo vecino. Cuanto mayor sea la masa del satélite, tanto más ancha será la zona, pero cuanto más grande sea la concentración de partículas en el anillo, tanto más frecuente será el que choques entre partículas envíen a algunas de ellas otra vez a la zona. Jack Lissauer y Frank H. Shu, de la Universidad de California en Berkeley, y M. Hénon, del Observatorio de Niza, han demostrado que si en los anillos de Saturno hubiera inmersos pequeños satélites de tamaño desde varios kilómetros hasta decenas de kilómetros, tales satélites podrían condicionar en buena medida la estructura fina de los anillos. Sin embargo, en el momento en que esto se escribe no se han detectado pequeños satélites, ni siquiera en las posiciones más probables, en las imágenes tomadas por los *Voyager 1* y *Voyager 2*.

Si el satélite está junto a un anillo, queda todavía otro efecto posible: las resonancias superpuestas que rodean la órbita del satélite pueden impedir que el anillo se expanda, y produzcan en él un borde bien definido. El borde exterior del anillo *A* de Saturno se mantiene así, probablemente, por intervención de los satélites 1980S26 y 1980S27. El anillo repele, a su vez, a los satélites. Además, ocurre que 1980S26 está muy próximo a una distancia radial de Saturno que lo encerraría en una resonancia con el satélite Mimas o Tethys, de masa mucho mayor. Tal resonancia "anclaría" su distancia radial, de modo que podría continuar esculpiendo el borde exterior del anillo *A*. La única dificultad que presenta esta hipótesis procede de unas mediciones precisas, que sitúan 1980S26 ligeramente fuera de la distancia de anclaje.

Por último, dos satélites cuyas órbitas estén próximas entre sí pueden impedir que un estrecho anillo entre sus órbitas se extienda radialmente. Por ejemplo 1980S26 y 1980S27, cuyas órbitas están separadas radialmente por sólo unos 2000 kilómetros, pueden actuar como pastores para poner en vereda o confinar el estrecho anillo *F*, compuesto de varias cintas, entre ellos. Cuando uno o ambos satélites pasan cerca de una cinta dada del anillo, su atracción

gravitatoria puede generar distorsiones, como las que se ven en las imágenes del anillo elaboradas por el *Voyager 1*. Sin embargo, todavía no está claro por qué el anillo *F* ha de constar de estrechas cintas múltiples, y no de una sola cinta, algo más ancha, ni por qué no está retorcida una de las cintas registradas por el *Voyager 1*. En imágenes hechas por el *Voyager 2*, el anillo *F* no presenta trenzas u otras distorsiones, y parece como si hubiera cambiado el número de cintas.

El que los pequeños satélites que actúan de pastores produzcan pequeños anillos de aspecto liso (así, anillos estrechos y excéntricos en el anillo *C* de Saturno y en la división de Cassini y probablemente en el anillo ε de Urano) o bien anillos enroscados o trenzados (los del anillo *F* de Saturno y los de la división de Encke, que las imágenes recabadas por el *Voyager 2* mostraron ser trenzados) puede depender del grado de opacidad de los pequeños anillos, o, en forma equivalente, de su densidad de partículas. En un anillo que sea bastante transparente, no suelen darse choques entre partículas, por cuya razón pueden persistir los "ecos" de perturbaciones gravitatorias antiguas. En un anillo que sea más bien opaco, la mayor frecuencia de los choques puede amortiguar las perturbaciones.

No disponemos todavía de un conocimiento cabal de la física de las interacciones entre pequeños satélites y anillos. Stanley F. Dermott y Thomas Gold, de la Universidad Cornell, han avanzado la hipótesis siguiente: un delgado anillo excéntrico, con bordes muy marcados, se conserva por la acción de un pequeño satélite único escondido en él y no por un par de satélites, grandes o pequeños, circundantes. El propio fracaso (de momento) en la búsqueda de pequeños satélites inmersos en los anillos principales de Saturno significa que hay que explorar otras explicaciones de la estructura fina radial en lugares tales como el anillo *B*. A este respecto se ha postulado una hipótesis que propone que las partículas del anillo obedecerían a aumentos, o disminuciones, transitorios de su densidad de agrupamiento con un ulterior aumento o disminución. Supóngase que la densidad de partículas aumenta localmente. Los choques entre las partículas se producen con mayor frecuencia y, como parte de la energía de los choques se pierde en calentamiento y deformación de las partículas, su energía para los movimientos aleatorios disminuye. A menos que los ulteriores choques a menores velocidades pierdan una proporción más pequeña de su energía, proseguirá el proceso de choque y soldadura. Se ha dicho que el resultado total podría dar lugar a variaciones radiales en la opacidad de un anillo.

Intervención de otras fuerzas

Las pequeñas partículas de los anillos pueden verse afectadas por fuerzas distintas de la de gravedad. Un importante ejemplo son las fuerzas electromagnéticas. Los anillos de Júpiter, de Saturno y posiblemente de Urano están en el interior de un plasma de baja densidad, esto es, se hallan dentro de un tenue gas formado por electrones cargados negativamente e iones cargados positivamente. Los electrones poseen una masa menor que la de los iones; se mueven, por tanto, a una velocidad mayor e inicialmente chocan con partículas del anillo más frecuentemente que los iones. Con el tiempo, las partículas llegan a cargarse negativamente por haber absorbido electrones. En ese punto, su carga repele la llegada de nuevas partículas del mismo signo. Lo que es más importante, las propias partículas del anillo se aceleran ahora al atravesar el campo magnético del planeta. Si las partículas no alcanzan a medir 0,1 micrometros, la fuerza electromagnética será mayor que la atracción gravitatoria del planeta y domina así el movimiento.

En ese marco admiten explicación algunos rasgos de la estructura de los sistemas de anillos. En el caso de Júpiter, por ejemplo, el eje del campo magnético está inclinado unos 10 grados con respecto al eje de rotación del planeta. Las fuerzas electromagnéticas, en tales circunstancias, pueden dar a las partículas pequeñas una componente vertical de velocidad mucho mayor que la de las partículas más grandes. La extensión vertical del halo del sistema de anillos de Júpiter es comparable con la que podría esperarse para partículas de 0,1 micrometros de tamaño o menores. En el caso de Saturno, las fuerzas electromagnéticas pueden resultar importantes para la peculiar estructura del anillo *F*. Se han propuesto, además, varias teorías ingeniosas para explicar cómo se originan las cuñas radiales en el anillo *B* de Saturno sin que se produzca una distorsión simultánea de la estructura radial, en forma de finas bandas, del anillo. Algunas de esas teorías recurren a una lluvia de partículas cargadas, procedentes bien del planeta o bien de los propios anillos, como manera de transferir carga eléctrica a partículas de tamaños del micrometro y levantarlas de la superficie de otras partículas mayores.

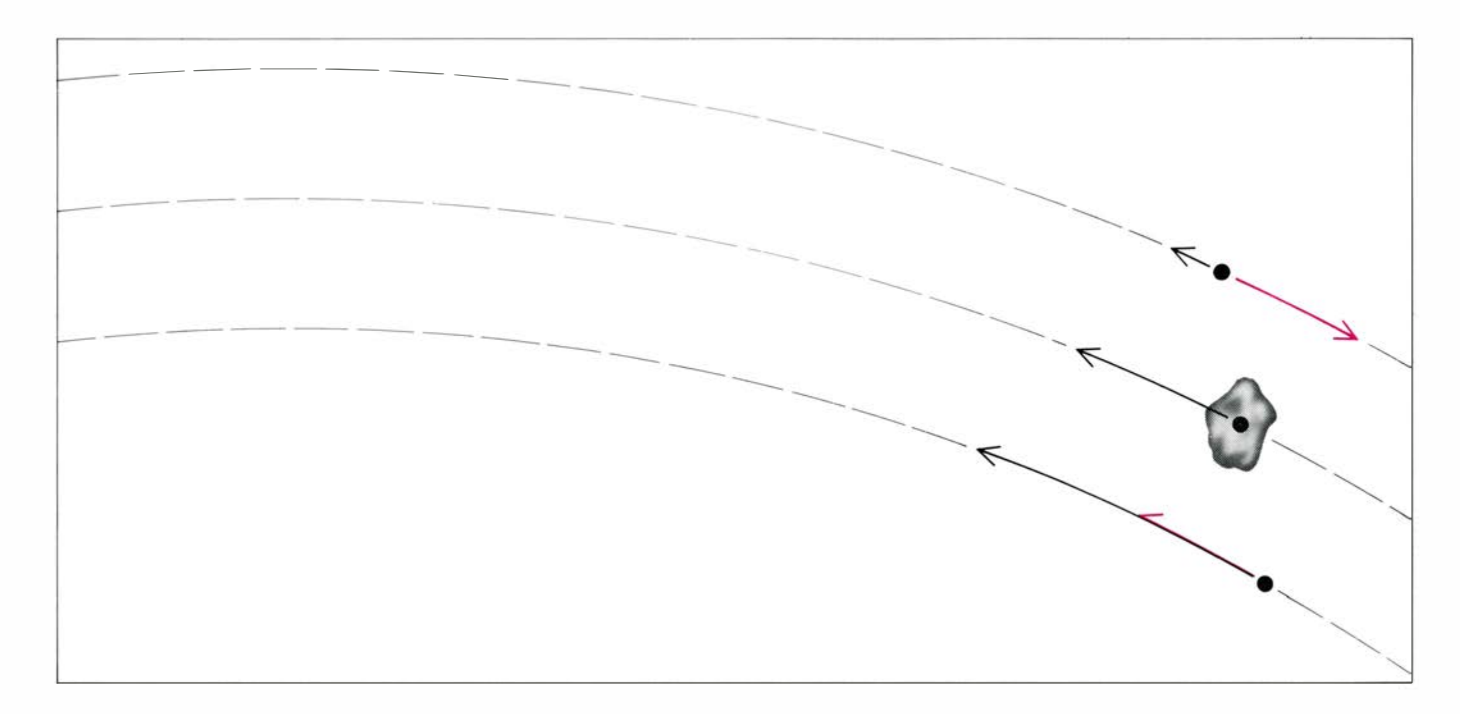
Una segunda fuerza que puede dominar el movimiento de pequeñas partículas en los anillos es la resistencia de los gases. Aquí el rozamiento debido a la presencia del plasma determina que las partículas de los anillos tiendan a caer sobre el planeta en trayectorias espirales. Cuanto menor sea la partícula, tanto más rápida será la reducción de su órbita. En sólo unos 20 años, por ejemplo, la resistencia del gas puede causar que una partícula de un micrometro de tamaño se mueva desde el borde exterior del anillo brillante de Júpiter hasta el borde interior. En otros 200 años cruzaría el disco difuso y quedaría dentro de la atmósfera del planeta. La apariencia relativamente lisa del anillo brillante de Júpiter puede deberse, en parte, a este rápido movimiento radial de las partículas pequeñas en su superficie. De hecho, el disco difuso puede estar poblado de partículas que han llegado a él procedentes del anillo brillante.

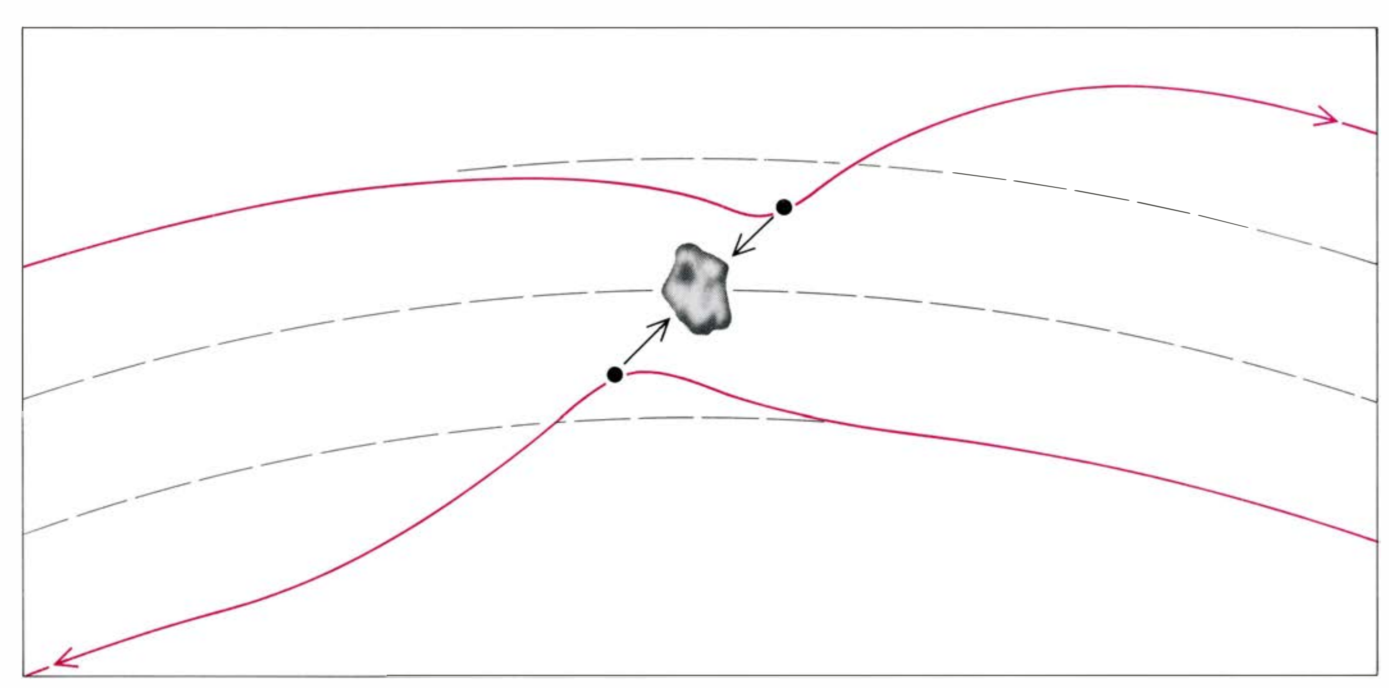
Tenemos la gravedad, las fuerzas electromagnéticas y la resistencia del gas. ¿Interviene algo más? Efectivamente. Las partículas pequeñas de un sistema de anillos están sometidas a choques que las destruyen. Por una parte, el espacio interplanetario contiene cuerpos sólidos cuyo tamaño varía desde menos de un micrometro hasta más de un kilómetro. Los cuerpos más pequeños son, con mucho, los que predominan en número. Los cuerpos interplanetarios están en órbita alrededor del Sol, mientras que la materia de los anillos gira alrededor del planeta. De ahí que ambas clases de cuerpos puedan chocar a velocidades de varias decenas de kilómetros por segundo. Si el cuerpo interplanetario midiera algo más de la centésima parte del tamaño de la partícula del anillo, el choque destruiría la partícula. Sobre esta base podemos calcular que una partícula de tamaño de un micrometro en cualquiera de los sistemas de anillos dura solamente unos 100.000 años. Hay otro tipo de choque que puede resultar más destructivo todavía. En los anillos de Júpiter y en los anillos exteriores de Saturno es probable que una partícula menor de 10 micrometros sea erosionada por choques con iones de alta energía en el cinturón de Van Allen del planeta, mucho antes de que un micrometeorito choque con ella.

¿Cómo se formaron los anillos? El Sol, los planetas, los asteroides y los cometas del sistema solar se formaron todos ellos, según se cree, hace unos 4600 millones de años dentro de la nebulosa solar, una nube difusa de polvo y gas. El Sol y los planetas gigantes, Júpiter, Saturno, Urano y Neptuno se formaron

en gran parte (y quizás enteramente) de los gases de la nebulosa; Mercurio, Venus, la Tierra, Marte, Plutón, los satélites, los asteroides y los cometas surgieron a partir de materia solidificada. La composición de ésta habría dependido de la temperatura de la nebulosa de un lugar a otro. En general, sin embargo, la materia solidificada constaba de diversas mezclas de "roca" (que incluía silicatos y hierro) y "hielos" (agua incluida).

Al principio, Júpiter y Saturno, y quizás Urano y Neptuno, superaban en varios cientos de veces sus dimensiones actuales. Bajo la influencia de sus propios campos gravitatorios se fueron contrayendo gradualmente. Al hacerlo, fueron girando más rápidamente. Con el tiempo, la creciente fuerza centrífuga asociada a la rotación hizo que la parte más externa de cada planeta gigante que iniciaba su formación –en esencia una envoltura gaseosa– divergiera de la concentración, más esférica, de gases en su parte interior; esta última continuó contrayéndose. La envoltura sirvió de fuente de los granos sólidos a partir de cuya acumulación se originaron los satélites grandes de cada planeta. Algu-





CONDUCCION GRAVITATORIA de partículas por pequeños satélites en el interior de un sistema de anillos. Ese "encarrilamiento" explica parte de la estructura en forma de bandas de los anillos de Saturno. El dibujo de arriba muestra un pequeño satélite y dos partículas representativas en órbita alrededor de un planeta. Según las leyes físicas, la partícula interior se mueve más rápidamente que el pequeño satélite, que, a su vez, lo hace más deprisa que la partícula exterior (*flechas negras*). Así, al pequeño satélite lo está adelantando la partícula interior, al tiempo que él rebasa a la partícula exterior (*flechas en color*). Cada partícula es atraída hacia el satélite pequeño por la acción gravitatoria de éste; de ahí que cada partícula quede un poco más cerca del pequeño satélite inmediatamente después de pasar a su misma altura que antes. El dibujo de abajo da cuenta del resultado. El tirón gravitatorio neto que el satélite ejerce sobre la partícula exterior tiene la dirección del movimiento orbital de esta partícula, que pasa, así, a una órbita más alta. A la inversa, el tirón que el pequeño satélite ejerce sobre la partícula interior se opone a la dirección del movimiento orbital, y esta partícula cae a una órbita más baja. En definitiva, el pequeño satélite limpia una banda a uno y otro lado de su trayectoria. Cuanto mayor sea la masa del satélite, más ancha será la banda.

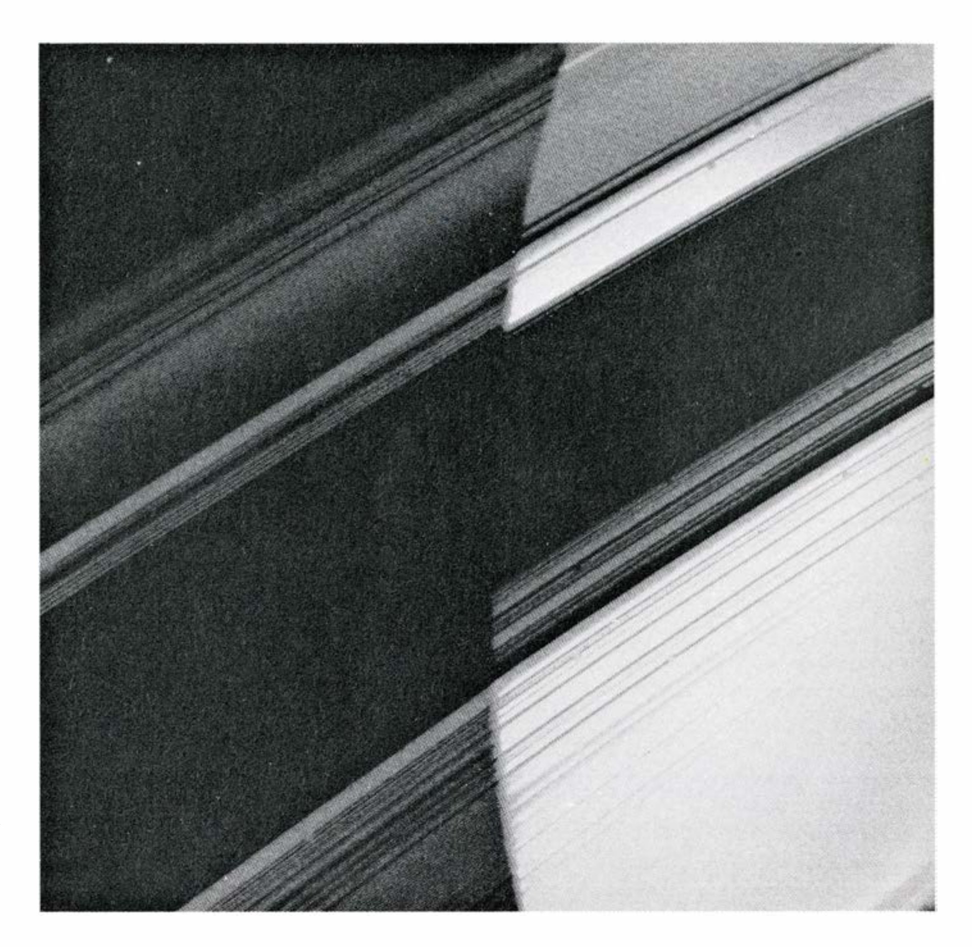
nos de los granos se formarían a raíz de la condensación de gases, vapor de agua por ejemplo, a medida que la envoltura se enfriaba. Quizá el mismo proceso de acumulación de granos para formar objetos de mayor tamaño dio lugar a los sistemas de anillos algo más cerca del planeta.

Tras un intervalo temporal del orden del millón de años, desaparecieron las envolturas circumplanetarias. Tal vez fueron arrastradas por un viento de gas ionizado procedente del joven Sol, o tal vez el rozamiento con las envolturas las hizo caer sobre sus planetas. En cualquier caso, así terminó la formación de satélites, pequeños satélites y anillos.

Condiciones circumplanetarias

Júpiter y Saturno envían al espacio en forma de radiación infrarroja aproximadamente el doble de la energía que absorben en forma de radiación solar. El exceso representa la conversión de energía gravitatoria en calor debido a la contracción pasada y presente de cada planeta. En la época en que se estaban formando los satélites, los planetas se contraían con mayor celeridad. En consecuencia, Júpiter, Saturno y demás planetas gigantes pueden haber radiado muchos miles de veces más calor que el que emiten en la actualidad. Este calor controló, con toda verosimilitud, la temperatura y, por ende, la composición de la materia sólida en las envolturas circumplanetarias. En un instante dado, la temperatura habría aumentado hacia el planeta, y, en un lugar determinado, la envoltura se habría ido enfriando continuamente a medida que transcurría el tiempo.

Hay varias teorías sobre el origen de los sistemas de anillos que existen alrededor de los tres planetas gaseosos gigantes, que postulan lo siguiente: el material de los anillos se formó en las envolturas circumplanetarias en posiciones próximas a sus actuales distancias radiales. En ese sentido, las variaciones de temperatura y otras condiciones en la vecindad de cada planeta constituyeron tal vez la causa de sorprendentes diferencias en la composición de sus satélites próximos y de la materia de los anillos. Las densidades medias de los cuatro satélites mayores de Júpiter implican que los dos más internos constan, exclusivamente, de rocas y los otros dos, de cantidades comparables de rocas y hielos. Puesto que el anillo de Júpiter queda muy dentro de la órbita del satélite grande más interno, las partículas que se formasen en la vecindad del anillo deben estar constituidas por sólo roca. Lo cierto es que únicamente pueden estar formadas por minerales bastante raros que se condensan a altas temperaturas.



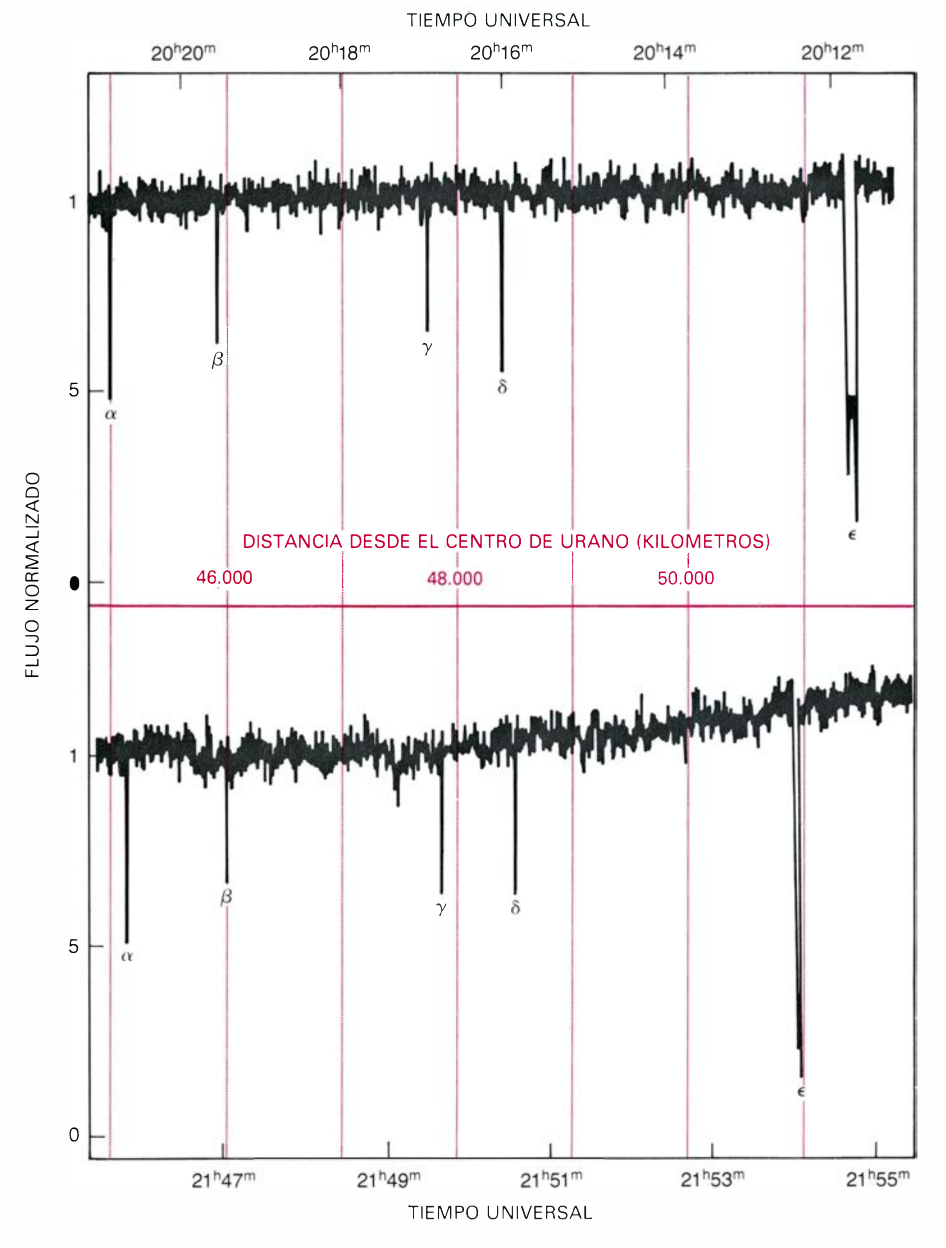
DIVERSOS GRADOS DE TRANSPARENCIA y, en consecuencia, de densidad de partículas por unidad de volumen, en los anillos de Saturno, puestos de manifiesto en esta imagen del *Voyager 1*. Se trata de la cara no iluminada de los anillos donde atraviesan el disco brillante del planeta. El disco brilla a través de la mayor parte del anillo *C* (*abajo*), y lo hace en grado similar a través de la división de Cassini, aunque ambos incluyen anillos delgados de opacidad mayor. Gran parte del anillo *B* (*centro*) es casi opaca.

Por contra, los satélites grandes más internos de Saturno constan, en esencia, de hielo. Así, las temperaturas, incluso a las pequeñas distancias que separan hoy al planeta de sus anillos, tienen que haber sido lo suficientemente bajas para permitir la formación de partículas que, en gran medida, son hielo auténtico, es decir, agua. Las diferencias entre las condiciones alrededor de Júpiter y las que predominaron alrededor de Saturno admiten una explicación fácil. En primer lugar, la cantidad de calor generado en los tiempos primitivos por la contracción de una esfera gaseosa depende aproximadamente del cuadrado de su masa. Júpiter, que multiplica por más de tres la masa de Saturno, emitió 10 veces más calor. Cabe también que no hubiera mucha roca disponible en la proximidad de Saturno para incorporarse en forma de partículas. Cuando se originó la envoltura circumplanetaria de Saturno, la atmósfera del planeta se extendía probablemente hasta regiones hoy ocupadas por los anillos y el satélite más interno. De aquí que la mayoría de los granos rocosos allí existentes puedan haberse conservado en la atmósfera de Saturno. O quizá se vieron atraídos por los satélites situados más hacia fuera.

Las hipótesis de que la temperatura en la vecindad de Saturno era relativamente baja y de que había poco material rocoso disponible concuerdan bien con el descubrimiento de que los satélites interiores de Saturno tienen una densidad media más baja, y por tanto albergan una mayor fracción de hielo, que la mayoría de los satélites exteriores. Estas hipótesis valen también para explicar por qué los anillos de Saturno parecen ser de hielo verdadero, casi puro, mientras que los de Júpiter son de roca. Se cree, sin embargo, que los anillos de Urano son de roca, cuando cabría esperar que lo fueran de hielo.

Hipótesis rivales

Las hipótesis avanzadas para dar cuenta de los anillos que ciñen a los planetas gaseosos gigantes discrepan, so-



LOS ANILLOS DE URANO se identificarón mediante una serie de atenuaciones registradas cuando una estrella pasaba por detrás del planeta, en 1977. La disposición de las atenuaciones, minutos antes de que la estrella desapareciera detrás del planeta (*arriba*), coincidía con la disposición minutos después de reaparecer (*abajo*). Para cuatro de las atenuaciones, la coincidencia era casi exacta; lo que significa que cuatro anillos son casi circulares. Un quinto, designado epsilon, ofrece una notable excentricidad. Observaciones posteriores han mostrado cuatro anillos más. Los datos fueron recogidos por James L. Elliot.

bre todo, a la hora de justificar la relación entre las partículas de los anillos que se observan hoy y la primitiva materia de los anillos. Según una hipótesis, un gran cuerpo único se fragmentó en miríadas de trozos cuando llegó a la proximidad de un planeta; y los fragmentos dieron origen después a los anillos. El cuerpo pudo haber sido un gran meteorito que sufrió un encuentro gravitatorio casual con el planeta, o quizás un pequeño satélite que se formó en la envoltura del planeta. En cualquiera de los dos casos, el agente de la fragmentación hubo de ser, a buen seguro, la distorsión por efecto de marea: la fuerza de cizalla que aparece por el hecho de que la atracción gravitatoria ejercida por el planeta sobre el cuerpo es mayor en las regiones del cuerpo más cercanas al planeta que en las que se hallan más alejadas. La creación de los anillos de Saturno por distorsión de marea fue propuesta, por primera vez, por el matemático francés Edouard Albert Roche en 1848. Roche había calculado que las fuerzas de marea excedían la autogravitación cohesiva de un satélite líquido, si éste se acercaba hasta una distancia algo menor de 1,5 radios de Saturno desde la superficie del planeta. Este umbral de rotura –límite de Roche– se encuentra próximo al borde exterior de sus principales anillos.

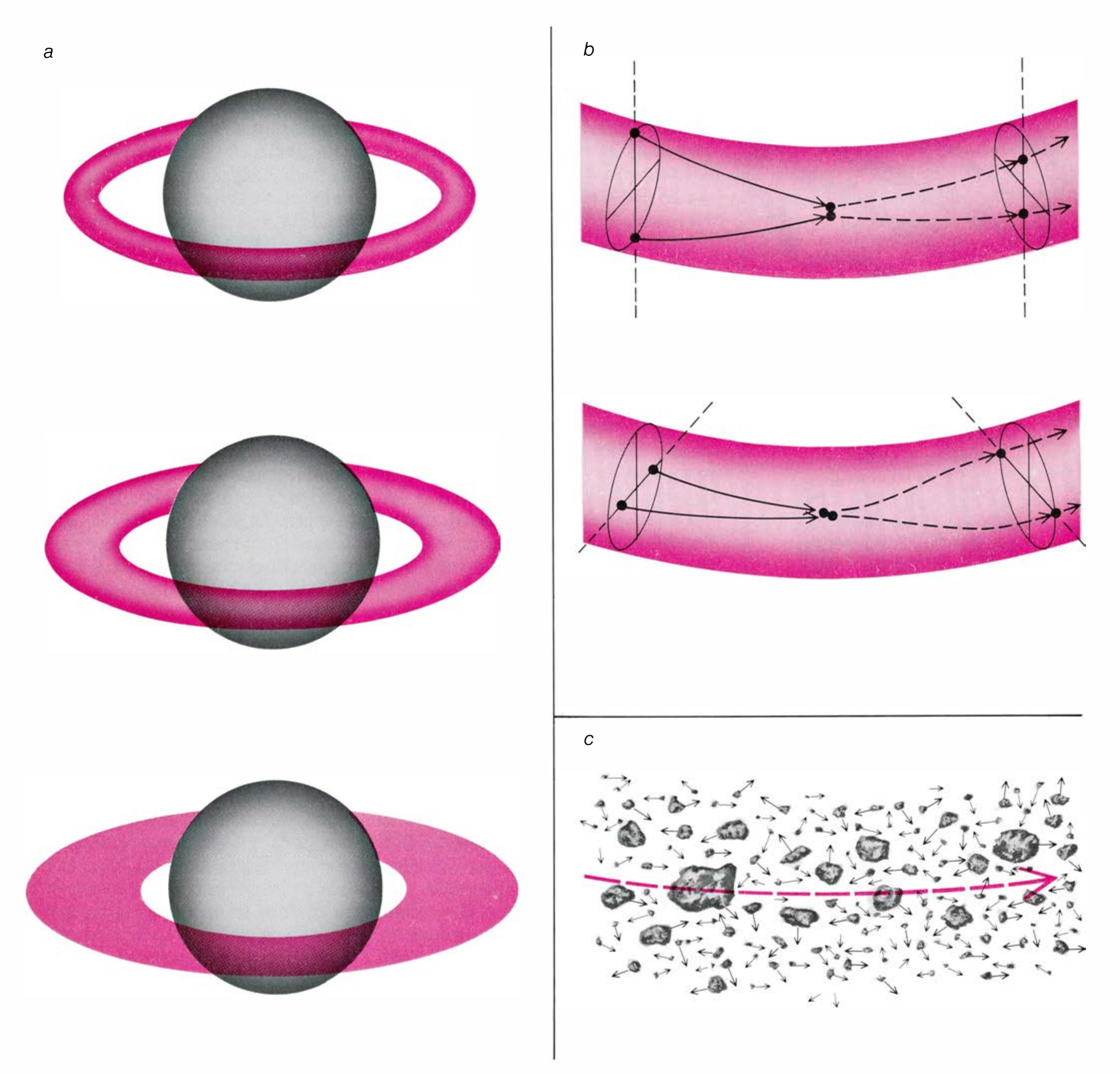
Resulta muy improbable, sin embargo, que un pequeño satélite próximo a Saturno fuera líquido. Tendría que haber sido sólido; ahora bien, un pequeño satélite sólido se mantiene unido no sólo por autogravitación sino también por las fuerzas que ordenan los átomos en la materia cristalina. Según Hans R. Aggarwal y Verne Oberbeck, del Centro de Investigación Ames de la NASA, las fuerzas de marea no pueden despedazar, a ninguna distancia de la superficie de un planeta, ningún pequeño satélite sólido de un diámetro algo menor de 100 kilómetros. Por si fuera poco, un satélite sólido mayor no puede despedazarse a distancia mayor de 0,4 radios planetarios de la superficie. Esta distancia coloca el umbral de destrucción dentro del borde interior de los principales anillos de Saturno. Es también improbable que el efecto de desintegración por marea de un meteorito errante proporcione partículas del anillo cerca de Saturno. Las partículas, como su cuerpo originario, tendrían velocidades suficientes para escapar de la vecindad del planeta.

Pero las fuerzas de marea pudieron intervenir de un modo más sutil, y no menos decisivo. Como ha demostrado Roman Smoluchowski, de la Universidad de Texas en Austin, la atracción gravitatoria de partículas de igual tamaño resulta insuficiente para mantenerlas juntas frente al efecto destructivo de las fuerzas de marea dentro del clásico límite de Roche (el límite calculado por Roche para un satélite líquido). Por el contrario, dos partículas que difieren grandemente en tamaño pueden resistir la ruptura a distancias en pleno interior del límite clásico. Ocurre, en realidad, que el límite impuesto para la ruptura por marea de dos partículas de igual tamaño y el límite impuesto para la destrucción del conjunto de dos partículas de tamaño desigual están próximos a los bordes exterior e interior de los sistemas de anillos que rodean a Júpiter, Saturno y Urano. Dentro del límite exterior de destrucción pudieron haberse acumulado partículas, pero tan lentamente que nunca formasen satélites grandes. Dentro del límite interior, el crecimiento resulta, quizá, poco menos que imposible.

La segunda hipótesis importante en relación con la historia de las partículas de los anillos fue propuesta por Eugene Shoemaker, del Servicio de Inspección Geológica de los Estados Unidos. Postula que un gran satélite, único, de la región de los anillos (o tal vez cierto número de satélites) chocaron catastróficamente con un meteorito errante. Las

fotografías de los satélites de Júpiter y Saturno hechas por los Voyager revelan, en efecto, que los satélites están marcados por un gran número de cráteres producidos por choques a alta velocidad. Mimas, uno de los satélites más pequeños de Saturno, tiene un cráter que abarca una tercera parte de un hemisferio. Faltó poco para que el choque que produjo este cráter destruyera a Mimas. Más cerca de Saturno hay dos objetos, de unos 100 kilómetros de diámetro, que ocupan órbitas casi idénticas. Quizá sean los fragmentos mayores de una colisión catastrófica entre un satélite y un meteorito gigante.

Hay varias razones por las que los choques catastróficos pudieron haberse producido preferentemente en la región que más tarde ocuparían los anillos. En primer lugar, los principales satélites de Júpiter y Saturno tienden a ser menores cuanto más cerca se hallan del planeta respectivo. Para una energía de choque dada, es más probable que se fragmenten los satélites pequeños que los grandes. En segundo lugar, el campo gravitatorio de un planeta focaliza la trayectoria de los meteoritos, de modo que el flujo de ellos que pasa cerca del planeta es significativamente mayor que el flujo a distancias crecientes.



EVOLUCION DE UN SISTEMA DE ANILLOS, iniciada, probablemente, en el primitivo sistema solar, cuando la contracción de una nube gaseosa en rotación para formar un planeta gigante como Saturno dejó tras sí una envoltura gaseosa extendida hacia fuera por la fuerza centrífuga asociada a su rápida rotación (*a*). Las partículas que se acumulaban en la envoltura deben haber estado en órbita alrededor del planeta, pero también se hallaban sujetas a movimientos aleatorios que las hacían chocar (*b*). Las colisiones eran inelásticas: deformaban las partículas y las calentaban. La correspondiente pérdida de energía cinética hacía decrecer su movimiento vertical. Los choques entre dos partículas a distancias radiales próximas tenían un efecto adicional. En cada choque, la partícula interna se movería más rápidamente en su órbita en torno al planeta y consiguientemente chocaría con la partícula más externa desde atrás. El intercambio de cantidad de movimiento elevaría la órbita de la partícula exterior y haría descender la de la interior. La envoltura se extendería horizontalmente. Por último, se formaría un disco plano donde las velocidades aleatorias relativas de las partículas serían sólo de una diezmillonésima parte de la magnitud de su velocidad orbital media (*c*). Las actuales partículas más pequeñas resultan de la erosión por choque de las partículas mayores.

Una última hipótesis en relación con la historia de las partículas de los anillos postula que los cuerpos mayores de los anillos vienen a ser mero resultado de la limitada extensión de la aglomeración de materia en la envoltura circumplanetaria a distancias próximas al planeta. La aglomeración comenzó con el enfriamiento de la envoltura y la condensación subsiguiente de materia gaseosa en minúsculos granos sólidos. Las fuerzas gravitatorias y el rozamiento gaseoso hicieron que los granos quedaran en el plano ecuatorial de la envoltura. Allí los granos continuaron creciendo por condensación del vapor en su superficie. Tal crecimiento podía conducirles hasta tamaños de algunos metros. Las partículas que forman la mayor parte de los anillos de Saturno varían en tamaño desde centímetros hasta metros. Cabe que resulten de ese proceso. Los pequeños satélites que posiblemente albergan los anillos en su interior serían el producto de una fase local de crecimiento en la que cuerpos del tamaño de metros quedarían soldados en virtud de choques suaves.

Las partículas de los tres sistemas de anillos son pequeñas y numerosas. Según la hipótesis de aglomeración, hay varias razones para que sea así. En primer lugar, la formación imperturbada de gránulos no podía empezar a una distancia dada de un planeta gigante en fase de gestación, mientras no se hubiera contraído hasta un tamaño inferior a esa distancia y se hubiera enfriado lo suficiente la envoltura circumplanetaria. Así, para la formación de gránulos en la proximidad de los planetas hubo menos tiempo disponible del que se disfrutó a mayores distancias. En segundo lugar, los gránulos de las proximidades de Júpiter se pudieron formar sólo a partir de las substancias relativamente raras y de alta temperatura de condensación que existían allí. Finalmente, el que un pequeño satélite alcanzase cierta magnitud significaría que sus resonancias superpuestas abarcarían una anchura comparable con su dimensión. No podría llegarle nueva materia; y, por tanto, dejaría de crecer. En los anillos de Saturno, el límite de crecimiento calculado está entre algunos kilómetros y algunas decenas de kilómetros. A mayores distancias de Saturno y de los otros planetas gigantes debieron de prevalecer condiciones diferentes, ya que allí se forman satélites mayores.

Erosión continua

¿Qué ocurre en los sistemas de anillos con las partículas pequeñas? Hemos indicado que el rozamiento gaseoso tiene por efecto el que las partículas de tamaño de un micrometro o menores describan espirales desde el borde exterior del anillo brillante hasta la atmósfera de Júpiter en escasos cientos de años. Manifiestamente, pues, tales partículas no pueden haber sobrevivido desde el tiempo en que el planeta tenía una envoltura gaseosa. Han de estar formándose hoy mismo. Según Joseph A. Burns, de Cornell, resultan de la erosión de cuerpos mayores en el anillo brillante de Júpiter, o en su proximidad.

No es fácil que las partículas que rondan el centímetro o lo sobrepasan ligeramente se destruyan por choque con micrometeoritos interplanetarios. Antes bien, cada choque excava un minúsculo cráter alrededor del punto de impacto y se expulsa una cantidad de materia 1000 a 10.000 veces mayor que la masa del cuerpo incidente. Muchas partículas del tamaño del micrometro de los sistemas de anillos pueden originarse, pues, como productos de eyección. Podemos suponer que si el diámetro de un satélite pequeño no llega a los 10 kilómetros, la mayoría de los productos eyectados a raíz de un choque con un cuerpo interplanetario escaparán del campo gravitatorio del pequeño satélite. Los productos eyectados que escapen del satélite carecerán, sin embargo, de la energía suficiente para escapar del planeta alrededor del cual se encuentra en órbita el pequeño satélite; describirían, pues, órbitas en el anillo. (La erosión incesante de una población originaria podría ser también fuente de los cuerpos localizados en los anillos de Saturno que miden del orden de un centímetro a un metro.)

En suma, los pequeños satélites y las partículas más grandes de los anillos datan probablemente de la primitiva historia del sistema solar: son contemporáneos de los satélites de los planetas gigantes. Por el contrario, las partículas más pequeñas se encuentran hoy en fase de formación. Se sospecha que los grandes satélites del sistema solar exterior, así como varios planetas (la Tierra incluida), surgieron por aglomeración de muchos cuerpos de tamaño menor. Seguramente, entre la muchedumbre de partículas de los anillos se reproducen hoy procesos similares, en menor escala. Los sistemas de anillos ofrecen, pues, un doble desafío. Hay que deducir los procesos que los formaron y, después, aprovechar ese conocimiento para poner de manifiesto cómo se constituyeron los satélites sólidos y los planetas.

Las fases fluidas de la materia

Más de un siglo costó describir la estructura de líquidos y gases, para volver hoy, de la mano de los ordenadores, a la vieja teoría molecular de las esferas macizas e inertes

J. A. Barker y Douglas Henderson

Bajo las condiciones de presión y temperatura que se dan normalmente en la superficie de la tierra, la diferencia existente entre un líquido y un gas salta a la vista. Los líquidos se pueden almacenar en vasijas abiertas y las rellenan totalmente hasta cierto nivel. Los gases no forman superficies libres, sino que tienden a ocupar todo el volumen disponible; deben, pues, guardarse en recipientes cerrados o mantenerse sometidos a un campo gravitacional, como sucede por ejemplo en el caso de la atmósfera de un planeta. Esta diferencia constituyó una de las características más destacadas de las primeras teorías elaboradas para describir las fases de la materia. Así, una de esas teorías sostenía, en el siglo XIX, que un líquido podía "disolverse" en un gas sin perder su identidad y otra afirmaba que las dos fases estaban constituidas por dos clases de moléculas diferentes: liquidones y gasones. Las teorías vigentes en la actualidad parten de un planteamiento distinto, prestando una especial atención a los rasgos comunes de los líquidos y los gases: ambos son formas de materia carentes de estructura permanente y ambos son susceptibles de fluir libremente. Tanto unos como otros son fluidos.

La semejanza fundamental que existe entre los líquidos y los gases se pone de manifiesto a partir de cierto valor de la temperatura y la presión. Supongamos que se calienta un recipiente cerrado, y que contiene determinada cantidad de líquido, aunque no lleno. El líquido se dilata, es decir, su densidad disminuye; una parte de él se evapora. En contraste, el vapor situado sobre la superficie libre del líquido se hace cada vez más denso a medida que nuevas moléculas evaporadas se le van incorporando. La dilatación del líquido y el aumento de la densidad del vapor aumentan la presión en el recipiente. Al seguir calentando, la densidad del líquido sigue bajando, y, subiendo, la del vapor hasta que llega un momento en que las dos densidades se igualan. La combinación de temperatura y presión a las que ambas densidades son iguales se conoce con el nombre de punto crítico. Por encima del punto crítico, el líquido y el gas ya no se pueden distinguir entre sí y sólo existe una fase fluida única de densidad uniforme.

El significado de las diferencias estructurales y dinámicas existentes entre los líquidos y los gases fue investigado, por primera vez, por el físico holandés J. D. van der Waals en la década de 1870. En aquel entonces se había desarrollado ya una teoría satisfactoria para los llamados gases ideales, según la cual la relación existente entre la temperatura, la presión y el volumen de un gas respondía a una ley particularmente simple. Van der Waals sugirió que todo líquido podía considerarse como un gas denso y, en consecuencia, describirse modificando adecuadamente la ecuación de estado de los gases ideales.

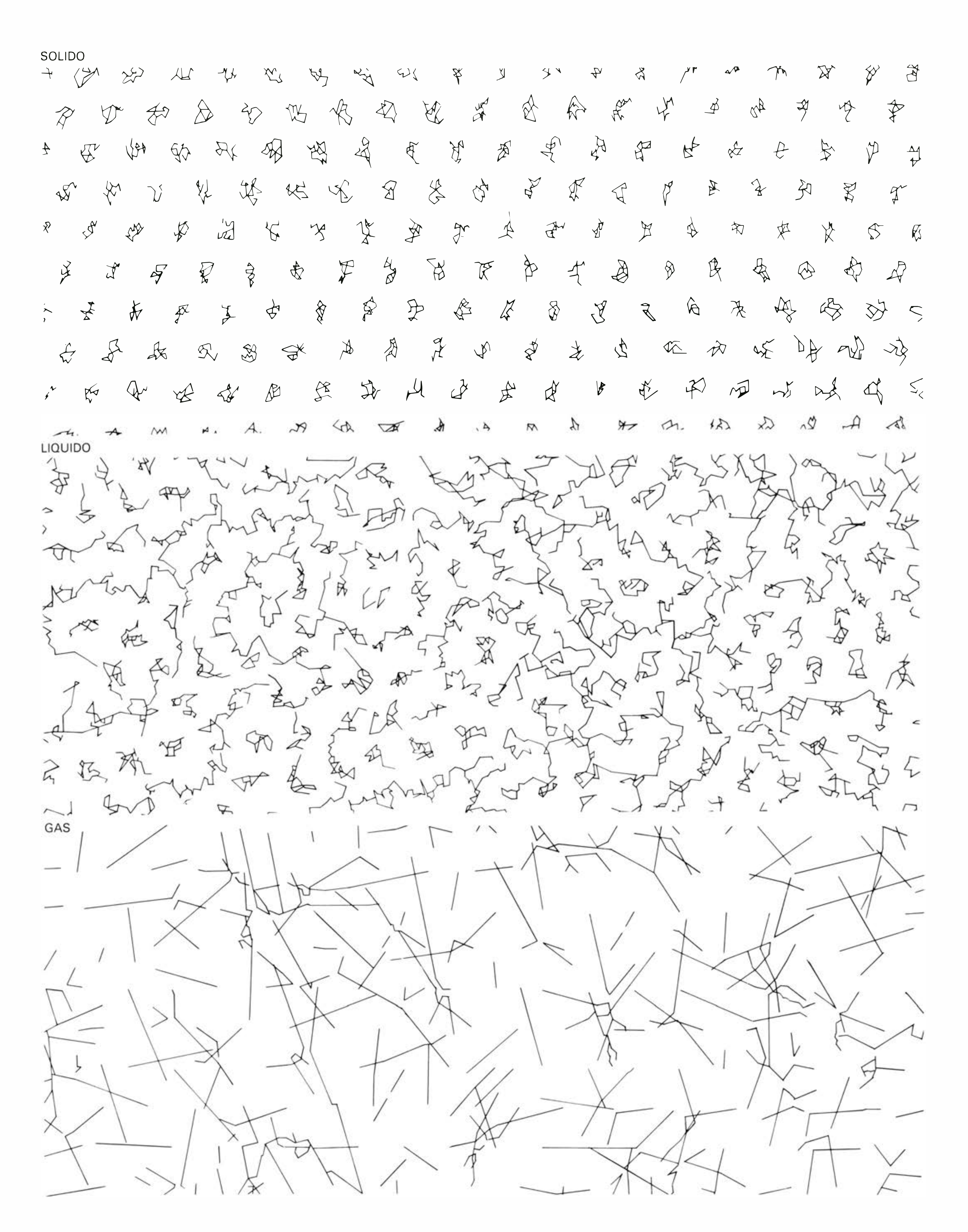
Las teorías más recientes de la estructura de los fluidos que vamos a estudiar a continuación tienen su origen en el trabajo de van der Waals. El modelo conceptual básico sobre el que se orientan se basa en suponer que los fluidos están constituidos por esferas macizas. Las propiedades de este modelo se pueden investigar mediante simulaciones realizadas con la ayuda de un ordenador. El modelo sirve también de punto de partida para describir los fluidos reales por métodos de perturbaciones, en las que el modelo básico se refina añadiéndole toda una secuencia de modificaciones de progresiva sutileza.

Prácticamente todos los elementos y compuestos químicos se presentan en las fases sólida, líquida y gaseosa. La transición de una fase a otra viene acompañada (o puede ser provocada) por un cambio en la temperatura, la presión, la densidad o el volumen. La naturaleza de tales transiciones se puede poner de manifiesto también tomando en consideración lo que sucede al calentar una sustancia en un recipiente cerrado, en una forma tal que permita controlar las diferentes condiciones que determinan su estado.

Supongamos que dicha sustancia se encuentra inicialmente en estado sólido y que se aumenta el volumen del recipiente manteniendo constante la temperatura. Al principio, se produce un aumento muy ligero en el volumen del sólido y, por tanto, disminuye la presión. Cuando el sólido empieza a fundirse, la presión se mantiene constante, debido a que el volumen total ocupado por el sólido y el líquido en coexistencia se compensa con el aumento de volumen del recipiente. Convertida toda la sustancia en líquido, la presión disminuye de nuevo, aunque no con la rapidez con que lo hacía cuando la sustancia era sólida. Tras un nuevo aumento en el volumen del recipiente se alcanza el punto de ebullición; este punto se caracteriza por una nueva estabilización de la presión, debido a que la formación de vapor compensa el aumento del volumen del recipiente. Finalmente, una vez se ha evaporado el líquido, la presión torna a reanudar su descenso.

Si realizáramos este mismo experimento a una temperatura superior al punto crítico de la sustancia, no se produciría una transición discontinua de líquido a vapor. Antes bien, la presión y la densidad del fluido menguarían de forma continua y regular. Hemos de resaltar, sin embargo, que la transición sólido-fluido se mantiene hasta las temperaturas más altas para las que las moléculas conservan su identidad. En otras palabras, no existe una temperatura crítica para el proceso de fusión.

El objetivo fundamental de una teoría que trate de describir las fases de la materia es dar cuenta de este tipo de observaciones en función de los movimientos de los distintos átomos o moléculas que integran la sustancia considerada. En otras palabras, se debe establecer una conexión entre las pro-



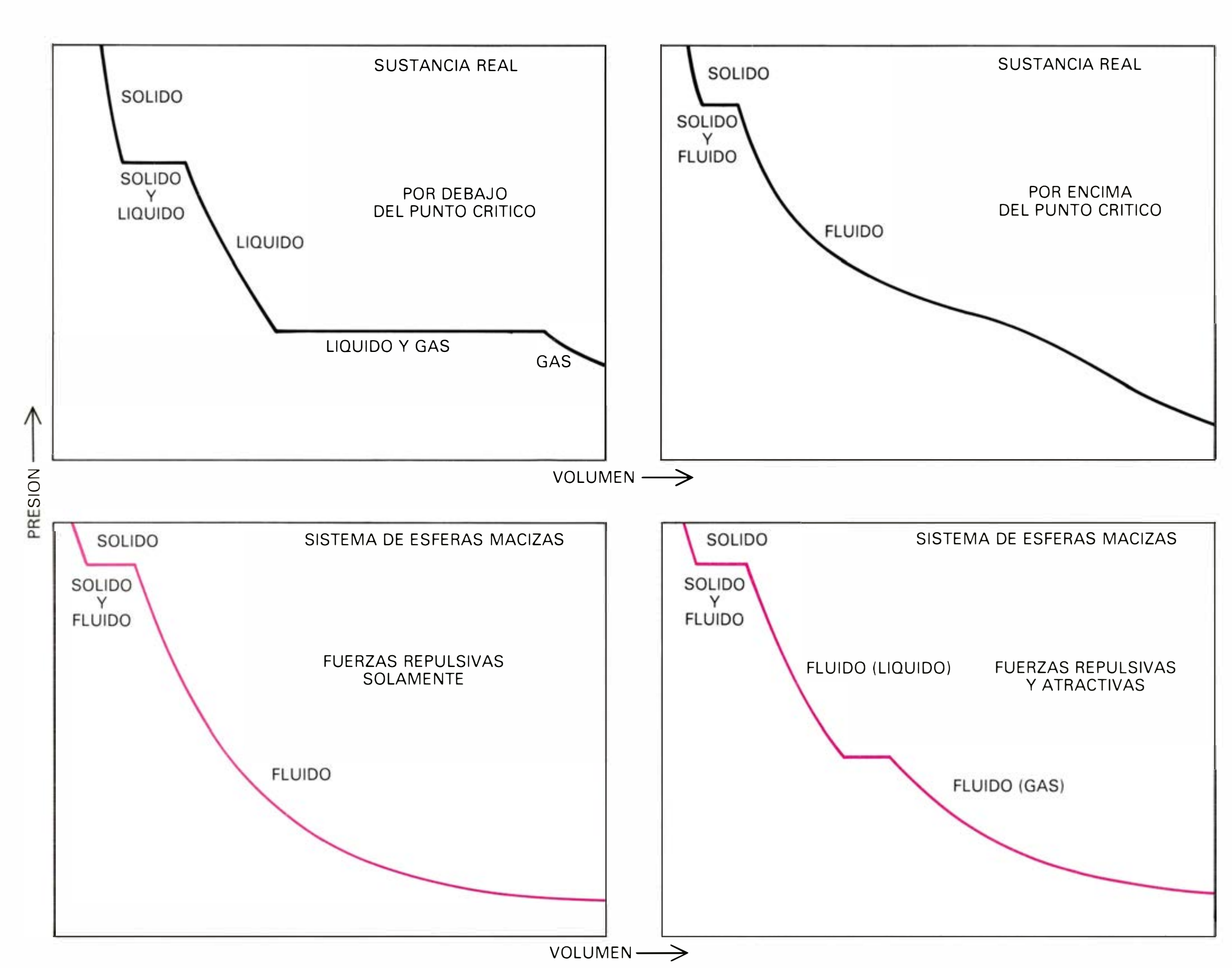
TRAYECTORIAS DE LAS MOLECULAS de un sólido, un líquido y un gas simuladas con la ayuda de un ordenador. El sistema de moléculas bidimensional considerado en estos cálculos tiene las mismas fases y las mismas transiciones de fase que la sustancia real, pero las posiciones de las moléculas y sus movimientos se ponen de manifiesto con mayor claridad que en ésta. En el caso de los sólidos, las moléculas se limitan a vibrar alrededor de sus posiciones fijas de la red cristalina, en tanto que las moléculas de los líquidos y de los gases se mueven libremente. Las únicas diferencias sustanciales entre los dos estados fluidos son la densidad y la frecuencia de las colisiones. El programa del ordenador calcula las trayectorias de las moléculas resolviendo las ecuaciones del movimiento, en dos dimensiones, de unas 500 moléculas. Esta simulación fue llevada a cabo por Farid Abraham, de IBM en San José, Cal.

piedades termodinámicas de las sustancias comunes y la mecánica de las moléculas individuales.

Bajo la mayoría de las circunstancias, los movimientos de las moléculas pueden describirse con aceptable precisión utilizando los métodos de la mecánica clásica o newtoniana. En mecánica clásica, la aceleración de una molécula (es decir, el cambio de su velocidad) resulta totalmente determinada por su masa y las fuerzas que actúan sobre ella. Si se conocen las posiciones y las velocidades de todas las moléculas de una cierta muestra de materia en un instante dado, y si se conoce la naturaleza de las fuerzas existentes entre las diferentes moléculas, la mecánica clásica permite calcular las posiciones y velocidades de esas moléculas en cualquier otro instante, tanto pasado como futuro. Llevando a cabo dicho cálculo, las transformaciones que constituyen los cambios de fase se pueden describir con todo detalle.

La mecánica clásica es incapaz de proporcionar una descripción adecuada de las fuerzas que se ponen de manifiesto entre las moléculas. Las únicas fuerzas significativas son de naturaleza electromagnética; estas fuerzas surgen como consecuencia de la distorsión que se produce en la nube de electrones que rodea una molécula cuando otra molécula se sitúa en sus proximidades. Dado que los electrones son mucho más ligeros que las moléculas y se mueven mucho más deprisa que éstas, la configuración de las nubes de electrones sólo se puede describir utilizando los métodos de la mecánica cuántica. Los cálculos basados en la mecánica cuántica encierran mayor dificultad que los basados en la mecánica clásica. Ahora bien, al definir las fuerzas intermoleculares se pueden introducir dos importantes simplificaciones. En primer lugar, precisamente por el hecho de que los electrones se mueven mucho más rápidamente de lo que lo hacen los núcleos atómicos, se puede suponer que los núcleos están en reposo. En segundo lugar, cabe suponer asimismo que la fuerza existente entre dos moléculas dadas es prácticamente independiente de la presencia de otras moléculas. En consecuencia, se puede deducir un cuadro bastante completo de cuantas fuerzas actúan en un sistema constituido por muchas moléculas (incluso en el caso de que estas moléculas estén dotadas de un movimiento rápido) a partir de la fuerza existente entre dos moléculas en reposo.

La tarea aparentemente sencilla de



PRESION Y VOLUMEN de una sustancia real y de un sistema modelo constituido por esferas macizas; ambas magnitudes varían fuertemente en los cambios de fase. Los gráficos superiores ponen de manifiesto la relación que existe entre la presión y el volumen de una sustancia real para dos temperaturas diferentes. Por debajo de la denominada temperatura crítica se observan dos transiciones de fase, cada una de las cuales viene caracterizada por una región plana en el diagrama presión-volumen. En estas regiones, la sustancia se separa en dos fases bien diferenciadas; una ligera variación en la presión o en la temperatura puede producir una transición de fase. Por encima de la temperatura crítica, la transición sólido-fluido persiste pero el líquido y el gas son indistinguibles entre sí; tan sólo existe una fase fluida supercrítica cuyo volumen puede variar de forma continua sin que tenga lugar un cambio de fase. Las gráficas de la parte inferior muestran la relación presión-volumen para un sistema de esferas macizas; en un caso actúan sólo fuerzas repulsivas y en el otro entran en juego, además, fuerzas atractivas. En el sistema de esferas macizas, la adición de fuerzas atractivas crea una zona plana muy parecida a la que separa al líquido del gas para temperaturas subcríticas en el fluido real. Las fuerzas atractivas son necesarias en la formación del estado líquido.

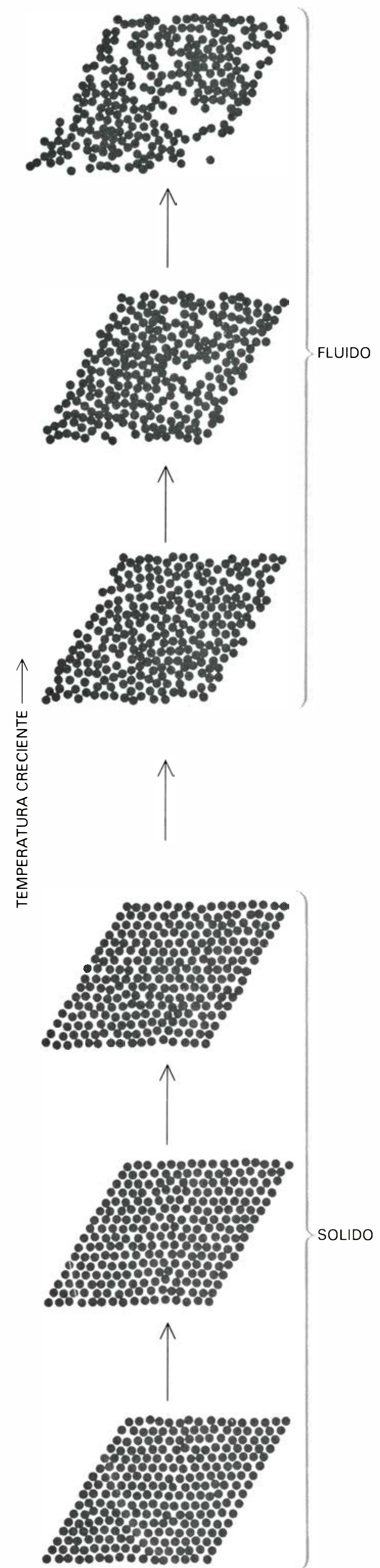
definir la naturaleza de la fuerza existente entre dos moléculas estacionarias resulta ser extraordinariamente difícil. De hecho, este problema sólo se ha podido resolver en el caso de los gases inertes, cuyas moléculas constan, en realidad, de átomos individuales cuya estructura electrónica es particularmente simple y poseen además simetría esférica. Esta última propiedad reviste una importancia enorme pues significa que la fuerza existente entre dos moléculas es la misma en todas las direcciones del espacio y depende tan sólo de la distancia que las separa. Las fuerzas que dependen también de la orientación de las moléculas pueden dar lugar a la aparición de propiedades muy interesantes en las fases condensadas de la materia; este es el caso, por ejemplo, del agua. Aun cuando el estudio de estas fuerzas anisotrópicas es un campo de investigación especialmente activo en la actualidad, vamos a centrar nuestra atención preferentemente sobre las fuerzas isotrópicas del tipo de las que se ponen de manifiesto en los gases inertes.

La fuerza existente entre dos átomos de un gas inerte se puede determinar de varias maneras. En el caso del helio, en el que el sistema constituido por los dos átomos incluye sólo cuatro electrones, las ecuaciones mecánico-cuánticas que definen la configuración de la nube de electrones se pueden resolver numéricamente con la ayuda de un ordenador potente. Estos cálculos han sido realizados por nuestros colegas Bowen Liu y A. Douglas McLean, del laboratorio de investigación de la International Business Machines Corporation en San José, California. Para gases inertes más pesados y para moléculas más complejas que éstas, la realización de tales cálculos no es factible; sin embargo, sí que cabe medir experimentalmente las fuerzas intermoleculares. Algunas de estas medidas son indirectas: el valor de dichas fuerzas se puede deducir a partir de ciertas propiedades de los gases que dependen de las colisiones moleculares, tales como la viscosidad o la conductividad térmica. Recientemente se han puesto a punto otros métodos de medida más directos, métodos en los que un haz de átomos se hace colisionar con otros átomos o en los que se analiza la radiación absorbida y emitida por moléculas diatómicas de gases inertes, cuya energía de enlace es muy baja.

Los resultados de estos cálculos o medidas suelen expresarse mediante una función de energía potencial, a partir de la cual se puede determinar fácilmente la fuerza existente entre dos moléculas. Esta función especifica la energía asociada con la interacción entre dos moléculas en relación con la distancia que las separa. Las moléculas tienden a aproximarse o alejarse entre sí, siempre que con ello disminuya su energía total; tendencia que se pone de manifiesto a través de una fuerza que empuja a las moléculas hasta que su separación sea la óptima. En un gráfico de la función de energía potencial, la fuerza intermolecular viene dada por la pendiente de la correspondiente curva cambiada de signo; en el punto en que la energía toma un valor mínimo la pendiente vale cero y, en consecuencia, dicha fuerza es nula. Cuando la separación es la óptima, la energía de la interacción es negativa, es decir, la energía del sistema constituido por las dos moléculas juntas es menor que la energía de esas mismas moléculas separadas.

En general, las funciones de energía potencial de casi todas las moléculas presentan básicamente la misma forma. Cuando las moléculas están infinitamente separadas no interaccionan en absoluto, y tanto la energía potencial como la fuerza que actúa entre las moléculas son iguales a cero. A medida que las moléculas se van aproximando mutuamente, sus nubes electrónicas se van distorsionando en una forma tal que la correspondiente energía total se reduce ligeramente; a raíz de lo cual surge una fuerza atractiva que tiende a aproximarlas todavía más. Este proceso continúa hasta que la separación entre dichas moléculas es del orden del diámetro molecular. Para acortar esa separación aún más, es necesario aportar energía desde el exterior, dado que se produce cierto solapamiento de las respectivas nubes electrónicas; en tales circunstancias, y como resultado de las interacciones repulsivas entre los electrones, la fuerza se hace intensamente repulsiva. La situación de equilibrio (energía mínima y fuerza nula) se alcanza cuando las interacciones atractivas y repulsivas son exactamente iguales.

La enorme sencillez de la función de energía potencial de un par de moléculas plantea de inmediato una cuestión importante: ¿en qué forma una in-



CONFIGURACION de las moléculas de una sustancia bidimensional puesta de manifiesto mediante "instantáneas" tomadas durante el proceso de simulación realizado por un ordenador. En el panel inferior la sustancia es un sólido, que se expande y se desordena ligeramente al aumentar la temperatura. Cuando se alcanza la temperatura de fusión, se produce una transición discontinua entre el sólido y el fluido. Las moléculas del fluido se mueven libremente, creando huecos transitorios a medida que la sustancia se expansiona. Esta simulación fue realizada también por Farid Abraham.

teracción tan simple como ésta puede dar lugar a las tres fases de la materia con esa extensa gama de propiedades diferentes? La función en cuestión sólo posee un punto de energía mínima. Las moléculas que constituyen un sólido se disponen en una forma tal que su separación resulta ser prácticamente igual que la distancia que corresponde al valor mínimo de la energía. ¿Cómo se explica entonces la existencia de los líquidos y los gases? Dado que su energía debe ser mayor que la que corresponde al mínimo, ¿por qué estas sustancias no reducen su energía convirtiéndose espontáneamente en sólidos? Cabría pensar que la función de energía potencial posee un segundo y un tercer mínimo para acomodar a las fases fluidas, pero la verdad es que no existen esos mínimos adicionales.

La resolución de esta paradoja se esconde en el hecho siguiente: el sistema de muchas moléculas que constituyen una determinada muestra macroscópica no es un sistema mecánico sencillo y la energía no es la única magnitud que influye en su comportamiento. El sistema de moléculas viene gobernado por las leyes de la termodinámica y la mecánica estadística, donde la temperatura y la entropía (o medida del desorden) intervienen también de una manera destacada.

Un sistema mecánico sencillo, pensemos en un balón rodando por la ladera de una montaña, alcanza el reposo cuando llega a su estado de energía más baja, es decir, cuando llega al fondo del valle. De forma análoga, dos moléculas aisladas tienden siempre a situarse a una distancia tal que su energía de separación sea mínima. Por otra parte, un sistema termodinámico no tiende al estado de menor energía, sino al estado de menor energía libre. La energía libre (A) es igual a la diferencia entre la energía (E) y el producto de la temperatura (T) por la entropía (S); en consecuencia, el comportamiento de un sistema termodinámico viene gobernado por la ecuación $A = E - TS$.

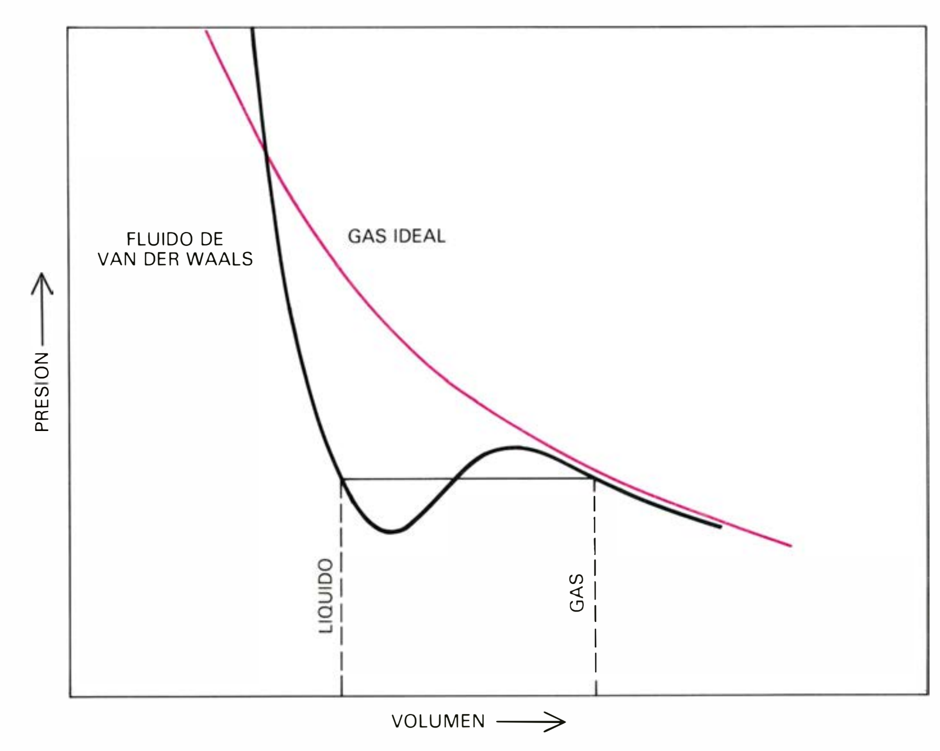
Esta relación permite deducir las características más fundamentales de un sistema termodinámico. A bajas temperaturas, la energía de las moléculas constituye la contribución más importante a la energía libre y, dado que esta energía tiene un valor mínimo para el estado sólido, dicho estado será el más favorecido. Cuando la temperatura es elevada, la entropía del sistema adquiere una importancia predominante. Como resultado, las fases fluidas son estables a altas temperaturas, incluso a pesar de que dichas fases corresponden a configuraciones cuya energía es mayor que la de la fase sólida.

Los conceptos de entropía y energía libre se pueden comprender fácilmente tomando en consideración la relación existente entre las descripciones macroscópicas y microscópicas de un sistema termodinámico. El estado macroscópico de un sistema viene especificado por magnitudes tales como la temperatura y la densidad, que son propiedades colectivas de todas las moléculas del sistema. El estado microscópico viene definido por la configuración detallada de sus mismas moléculas. Cabe destacar especialmente el hecho de que a un estado macroscópico dado pueden corresponder muchos estados microscópicos diferentes. En otras palabras, varias configuraciones diferentes de moléculas pueden dar lugar a los mismos valores de temperatura, densidad y otras magnitudes termodinámicas.

En mecánica estadística la entropía se interpreta como una medida de la probabilidad intrínseca de un determinado estado macroscópico. La probabilidad viene determinada por el número de estados microscópicos que corresponden a dicho estado macroscópico y por la probabilidad asignada a cada uno de esos estados microscópicos. Todos los estados microscópicos de igual energía tienen la misma probabilidad, pero la probabilidad de un estado disminuye fuertemente (siguiendo una curva exponencial) a medida que la energía de ese estado aumenta o que la temperatura disminuye. La expresión matemática que define la dependencia de la probabilidad con la energía y la temperatura se conoce con el nombre de factor de Boltzmann, en honor del físico austriaco Ludwig Boltzmann.

El factor de Boltzmann establece que los estados microscópicos individuales cuya energía es mínima son siempre los más probables, pero ello no quiere decir que el estado macroscópico observado tenga necesariamente una energía baja. Algunos estados macroscópicos de energía elevada pueden estar originados por un número mucho mayor de estados microscópicos de alta energía. Aunque la probabilidad que corresponde a cada uno de los estados microscópicos de alta energía sea baja, dichos estados pueden ser tan numerosos que la probabilidad de que el sistema ocupe uno u otro de ellos se torne muy elevada. Desde este punto de vista, podemos afirmar que los fluidos existen porque hay muchas más formas de organizar las moléculas en un líquido que en un sólido y todavía más en un gas.

La teoría de la mecánica estadística se presentó, por primera vez y en forma completa, en una obra del físico norteamericano Josiah Willard Gibbs: *Principios básicos de Mecánica Estadística*, publicada en el año 1902. Este autor demostró que un sistema de partículas



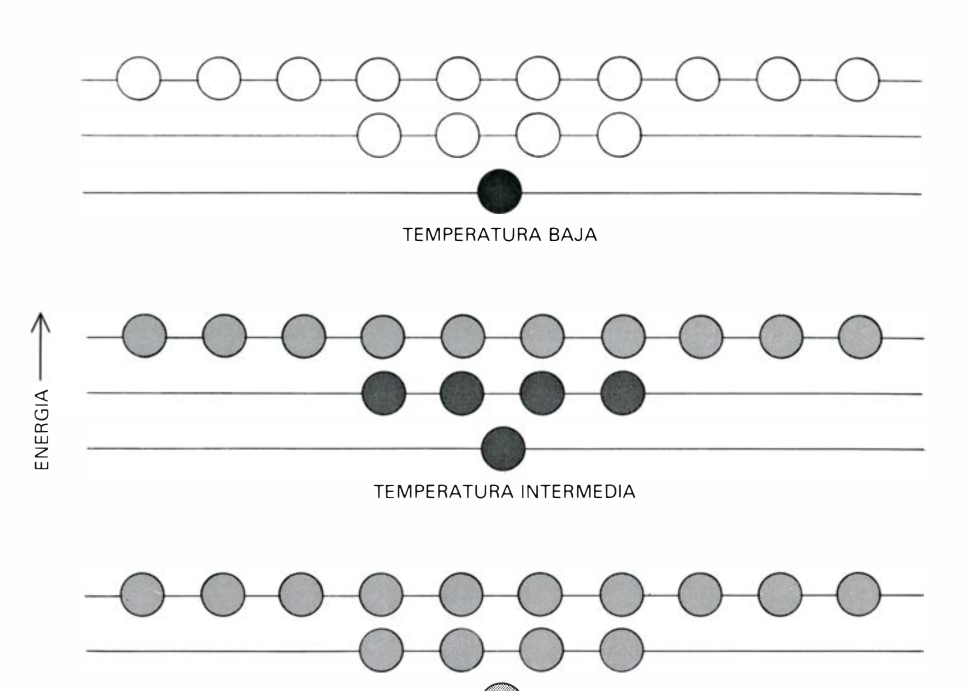
TEORIA DE LOS LIQUIDOS formulada por el físico J. D. van der Waals modificando adecuadamente la ecuación de estado que define la relación entre la presión, el volumen y la temperatura en un gas ideal. La depresión en el potencial de van der Waals es una consecuencia de la introducción de fuerzas atractivas entre las moléculas y de la modificación del volumen disponible, excluyendo el volumen que ocupan las propias moléculas. Las fuerzas atractivas intermoleculares reducen la presión total del fluido.

a una temperatura dada se distribuye por sí mismo entre todos los niveles energéticos disponibles. Los elegantes resultados de esta teoría llenaron el vacío que había existido hasta entonces entre las dos grandes concepciones del mundo: la mecánica y la termodinámica.

La comprensión de las fuerzas intermoleculares y la mecánica estadística posibilitó el manejo de los elementos básicos que se necesitaban para investigar los estados de agregación de la materia. Pero faltaban todavía ciertas herramientas de tipo práctico. A mediados del siglo XX existían un conjunto de teorías satisfactorias de los sólidos, sistemas altamente ordenados, y de los gases no demasiado densos, sistemas altamente desordenados. No existía, empero, una teoría adecuada para los líquidos y los gases comprimidos.

El principal impedimento para el desarrollo de esta teoría consistía en el hecho de que, en cualquier descripción de un líquido o un gas denso, debe tomarse en consideración un número muy elevado de moléculas. Estos problemas de muchos cuerpos no se pueden resolver analíticamente; lo cierto es que ni siquiera el problema de tres cuerpos admite una solución exacta, salvo en condiciones muy especiales. Para describir los fluidos densos es necesario desarrollar ciertos métodos aproximados. El método más directo consiste en la simulación del fluido con la ayuda de un ordenador.

El primer paso en la construcción de un modelo matemático de un fluido estriba en la elección de una función de energía potencial que sea capaz de dar debida cuenta de la fuerza existente entre cada par de moléculas. La función a utilizar puede calcularse teóricamente de acuerdo con la mecánica cuántica o deducirse por vía experimental; en cualquier caso, lo mejor es disponer de una función cuya estructura matemática sea lo más simple posible. Así, alrededor del año 1930, John E. Lennard-Jones introdujo una función sencilla que se aproximaba bastante a la forma de la función que determina la interacción existente entre los átomos de los gases inertes. Posteriormente se ha llevado a cabo una amplia investigación con una función todavía más simple que ésta: la que postula el modelo de la esfera maciza. En el tipo de fluido que considera este modelo no existe un potencial atractivo y las esferas son totalmente rígidas, hasta el punto de que no hay posibilidad alguna de interpenetración. En consecuencia, el potencial repulsivo se hace infinito para una determinada distancia de separación.



LA CONFIGURACION MICROSCOPICA de un sistema de moléculas determina el estado macroscópico de dicho sistema. Se representa aquí el número relativo de posibles configuraciones dibujando unos círculos sobre cada uno de los tres niveles energéticos considerados en el ejemplo. La probabilidad de que el sistema posea cierta configuración viene dada por la intensidad del sombreado del círculo. Esta probabilidad depende de la energía de la configuración y de la temperatura del sistema. Para bajas temperaturas, la configuración de menor energía es la única cuya probabilidad es apreciable; esta configuración corresponde al estado sólido. Para temperaturas intermedias, las configuraciones de energía baja resultan todavía favorecidas, pero dado que existen muchas más configuraciones de alta energía, cada nivel energético tiene aproximadamente la misma probabilidad total; esta situación prevalece durante la transición de sólido a la fase líquida. Cuando la temperatura es elevada, la probabilidad de todas las configuraciones es prácticamente la misma. Se tiene entonces la firme certeza de observar una configuración de alta energía, debido a que estas configuraciones son, ahora, las más numerosas. El sólido se ha fundido.

Un modelo se constituye por un conjunto de moléculas, a cada una de las cuales se le asigna una posición inicial. En cada etapa de la evolución del modelo se deben especificar las nuevas posiciones que ocupan dichas moléculas, posiciones que representan un nuevo estado del sistema. En algunos casos, las velocidades de todas las moléculas vienen también especificadas en cada etapa. El vapor promedio de todas las configuraciones observadas a lo largo de muchas etapas permite predecir las propiedades macroscópicas del modelo de fluido; posteriormente, se pueden comparar estas predicciones teóricas con los datos experimentales correspondientes a fluidos reales.

Los sucesivos estados del sistema considerado por este modelo se calculan mediante el método llamado de Monte Carlo. Este nombre responde al hecho de que el mecanismo esencial de este método consiste en un generador de números al azar cuya función es conceptualmente equivalente a la de una ruleta. El algoritmo en que se funda fue inventado en 1953 por Nicholas C. Metropolis, Arianna W. y Marshall N. Rosenbluth y Augusta H. y Edward Teller; fue utilizado, en su fase inicial, por William W. Wood. El algoritmo genera una secuencia de configuraciones de un sistema de muchas partículas que imita algunas de las propiedades estadísticas del movimiento molecular. La probabilidad de que una configuración dada aparezca en la secuencia es proporcional al valor del factor de Boltzmann para esa configuración en particular, exactamente igual a como sucede en un sistema termodinámico real.

En el método de Monte Carlo, el programa del ordenador mueve las moléculas de una en una de acuerdo con las leyes de la mecánica estadística. Sin embargo, debido al empaquetamiento compacto de las moléculas, un movimiento arbitrario puede llevar a una molécula en particular lo suficientemente cerca de otra como para que se produzca un cierto solapamiento de ambas. En este caso, las fuerzas repulsivas determinan que la energía potencial alcance un valor grande y positivo. Puesto que el factor de Boltzmann disminuye exponencialmente a medida que la energía aumenta, el citado solapamiento da lugar a que el factor de Boltzmann sea realmente igual a cero

para la mayoría de las configuraciones. En la práctica esta dificultad se supera calculando el cambio de energía que se produce en cada movimiento para cada configuración. Si una molécula se aproxima demasiado a sus vecinas, el programa la devuelve automáticamente a su posición original. Con esta modificación, la frecuencia de cada configuración es proporcional a su factor de Boltzmann.

B. J. Alder y T. E. Wainwright, del Laboratorio Lawrence de Berkeley, Aneesur Rahman, del Laboratorio Nacional de Argonne, y Loup Verlet, del Laboratorio de Física Teórica y de Altas Energías de Orsay, han puesto a punto otro método de simulación por ordenador. Este método se conoce con el nombre de dinámica molecular y se basa en resolver directamente las ecuaciones clásicas del movimiento de un sistema de moléculas. En cada etapa de la evolución del modelo, la función de energía potencial se emplea para calcular la fuerza que sobre cada una de las moléculas de un conjunto dado ejercen las moléculas próximas. Una vez conocido el valor de esta fuerza neta, se puede determinar la correspondiente aceleración y, a partir de ésta, calcular la nueva posición y la velocidad de la molécula considerada.

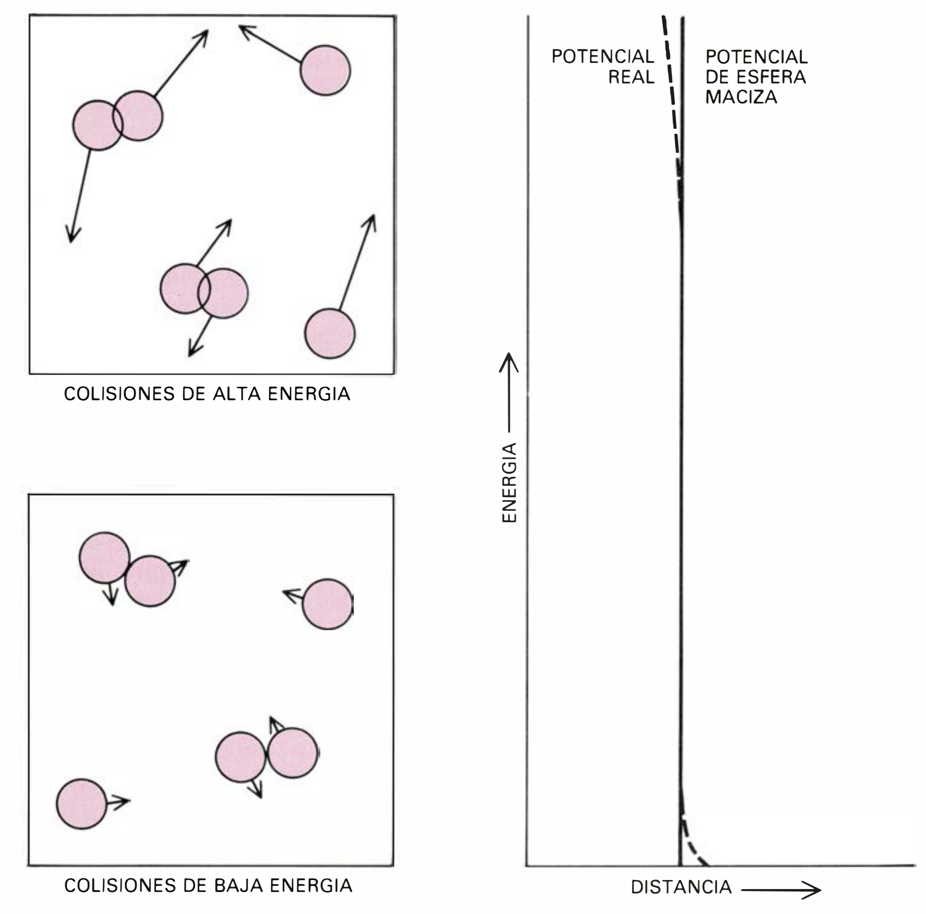
Contrariamente a lo que sucede con el método de Monte Carlo, el método de la dinámica molecular proporciona cierta información acerca de la evolución del sistema con el tiempo. En consecuencia, este método permite calcular el valor de algunas propiedades del sistema fluido que dependen del tiempo, como es el caso, por ejemplo, de la viscosidad. Otras propiedades, así la entropía, pueden conocerse también promediando las configuraciones observadas a lo largo de cierto período de tiempo.

La simulación por ordenador es capaz de proporcionarnos, además, una descripción del fluido conocida como función de distribución radial. Esta función nos da a conocer la forma en que la densidad media de las moléculas varía al alejarse radialmente de una cierta molécula de referencia. Utilizando un modelo del fluido existe un procedimiento inmediato para calcular la función de distribución radial. Este procedimiento consiste sencillamente en tabular las distancias que separan a una molécula determinada de todas las demás. La función de distribución calculada para este modelo se puede comparar entonces con la función que corresponde a un sólido o a un fluido real. La función de distribución radial de un material real se puede determinar experimentalmente bombardeando los átomos o moléculas en estudio con haces de rayos X o neutrones y midiendo los ángulos de difusión correspondientes.

La función de distribución radial permite diferenciar las distintas fases de la materia. Para distancias pequeñas, la función vale siempre cero, debido al hecho de que dos moléculas no pueden ocupar la misma posición. Para un gas a baja densidad y alta temperatura, la función tiene un valor prácticamente constante para cualquier distancia que sea superior a dicha zona prohibida, por la sencilla razón de que las moléculas se distribuyen al azar. En un sólido cristalino, la función de distribución presenta una serie de máximos muy acusados.

Las funciones de distribución radial más interesantes son las de los líquidos y los gases densos. Para una distancia del orden de un diámetro molecular, se presenta un primer máximo que corresponde a las moléculas más próximas. Por ser muy probable que las moléculas cercanas estén en contacto, las vecinas más próximas se situarán a distancias relativamente pequeñas y el máximo en cuestión será muy acusado. Tras este primer máximo, la función presenta un mínimo, seguido por un nuevo máximo menor que el anterior para una distancia del orden de dos diámetros moleculares; este segundo máximo corresponde a la formación de una segunda capa de moléculas. Para separaciones todavía mayores, los máximos se hacen más anchos; cuando se sobrepasan los cuatro diámetros, la modulación de la función de distribución se amortigua totalmente; en ese punto la densidad de las moléculas se aproxima a la densidad uniforme del líquido considerado como un todo.

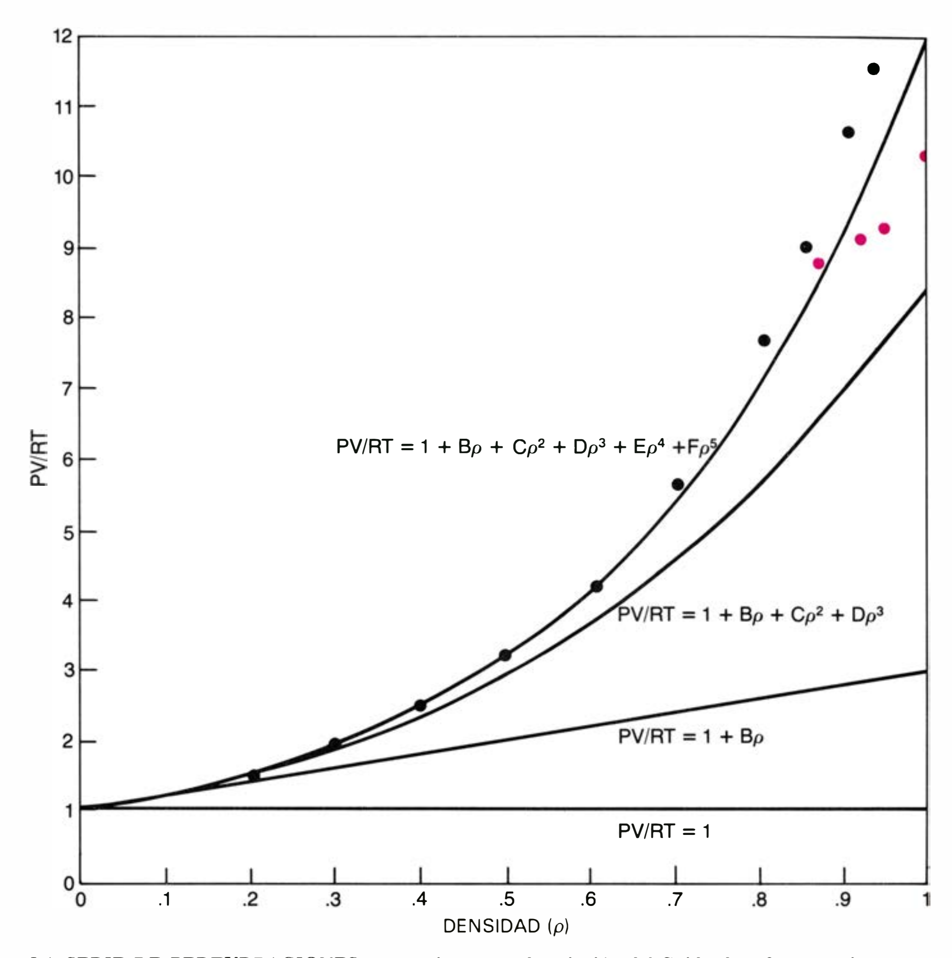
La función de distribución radial del argón se ha determinado con gran precisión por difusión de neutrones. Los resultados obtenidos en simulaciones con ordenador para el argón líquido están en excelente acuerdo con los datos experimentales correspondientes a la función de distribución radial y a las propiedades termodinámicas tales como la presión, el volumen y la temperatura.



A MEDIDA QUE AUMENTA LA ENERGIA o la temperatura, disminuye el diámetro efectivo de las esferas macizas. Cuando la energía es baja, las colisiones moleculares son por término medio menos fuertes y existe un solapamiento despreciable de las nubes electrónicas. Para energías altas, las colisiones son más violentas y, mayor, el solapamiento de las nubes electrónicas. En este caso, los centros de las moléculas se aproximan entre sí mucho más; su diámetro eficaz resultará menor. Este efecto se introduce en un fluido de esferas macizas postulando que el diámetro molecular es función de la temperatura.

Aunque los métodos de Monte Carlo y de la dinámica molecular utilizan algoritmos diferentes, los detalles operativos son similares. Si se desea llevar a cabo los cálculos pertinentes en un tiempo razonable, el número de moléculas a considerar debe ser limitado. En la mayoría de las aplicaciones, este número es del orden de 100 o 1000. En general, se supone que las moléculas llenan un recinto cuyas dimensiones vienen determinadas por la densidad del fluido. Si el número de moléculas contenidas en el recinto es pequeño, los efectos de superficie adquieren una importancia desmesurada.

El programa del ordenador corrige en parte los efectos de superficie utilizando condiciones de contorno periódicas. El recinto está rodeado por todos lados por muchas réplicas periódicas de sí mismo, como las células unitarias de un cristal. Cada una de las moléculas encerradas en el recinto puede interaccionar con las moléculas "fantasma" que existen en los recintos réplica. Por otro lado, cada vez que una molécula sale del recinto, su doble entra automáticamente por el lado opuesto. De esta forma, el número de moléculas a considerar se mantiene constante.



LA SERIE DE PERTURBACIONES proporciona una descripción del fluido de esferas macizas cuya precisión mejora sensiblemente conforme crece el número de términos de la serie. Dicha serie se inicia con el gas ideal (es decir, un sistema en el que las moléculas no interaccionan entre sí). Podemos formalizar la ecuación de estado del gas ideal del siguiente modo: $PV/RT = 1$, donde P representa la presión, V el volumen, R una constante y T la temperatura. De aquí que para el gas ideal el valor de PV/RT permanece constante e igual a la unidad, cualquiera que sea su densidad. Para un gas real a temperaturas elevadas, el valor de PV/RT tiende a aumentar con la densidad. La teoría da cuenta de este aumento añadiendo a la ecuación de estado una serie de términos en los que la densidad aparece elevada a potencias cada vez mayores. En consecuencia, la ecuación toma la forma $PV/RT = 1 + B\rho + C\rho^2 + D\rho^3$, y así sucesivamente. La suma de seis de estos términos suministra una buena aproximación a los resultados obtenidos en los cálculos basados en el modelo de esferas macizas. Estos cálculos han sido llevados a cabo por los autores de este artículo (*puntos negros*) y por B. J. Alder y T. E. Wainwright (*puntos en color*). Sin embargo, el gas ideal no proporciona un sistema de referencia satisfactorio para los modelos del estado líquido, donde el valor de los términos sucesivos de la serie no disminuye progresivamente, sino todo lo contrario.

Se ha demostrado que un sistema constituido tan sólo por 100 moléculas puede proporcionar unas estimaciones muy precisas de las propiedades de las fases condensadas de sustancias sencillas, tales como el argón. La única región del diagrama fásico en la que los modelos fallan se sitúa en las proximidades del punto crítico. Cerca del punto crítico se producen variaciones de la densidad y de otras propiedades a lo largo de muchas escalas de tamaño. Las interacciones de largo alcance, entre las que se incluye la interacción electrostática entre los iones y la interacción entre moléculas polares tales como las del agua, presentan asimismo dificultades especiales.

Las simulaciones llevadas a cabo con un ordenador digital suponen un intento de calcular las propiedades de un fluido a partir de principios básicos. La segunda línea fundamental en el desarrollo de la teoría de los fluidos parte de un planteamiento diferente. En primer lugar, se elabora un modelo matemático preliminar, capaz de proporcionar una descripción razonable del sistema, y, a continuación, se introducen en dicho modelo toda una serie de refinamientos sucesivos.

Este tipo de procesos se conoce con el nombre de métodos de perturbaciones; fueron utilizados por primera vez en el cálculo de órbitas planetarias. En ese contexto, la aproximación inicial consiste en considerar que el Sol y el planeta en cuestión están aislados. A continuación se introducen las perturbaciones que sobre la órbita calculada en primer término produce la presencia de los otros planetas. Estas correcciones suelen tomar la forma de una serie de potencias en las que una pequeña cantidad x aparece elevada a potencias cada vez mayores. Cada uno de los términos de la serie describe el efecto de una perturbación diferente. La evaluación de la serie se hace más y más difícil a medida que se añaden nuevos términos; mas, si x es pequeña, la importancia de los sucesivos términos disminuye rápidamente, dado que dicha cantidad está elevada a potencias grandes. En consecuencia, basta con considerar unos pocos términos para alcanzar la precisión deseada. El problema más difícil de este método radica precisamente en encontrar una cantidad x que sea lo suficientemente pequeña como para proporcionar unos resultados suficientemente precisos utilizando un número razonable de términos.

Al aplicar los métodos de perturbaciones a la teoría de los fluidos existe un candidato muy claro como sistema de referencia: el gas ideal, es decir, un sistema constituido por moléculas que no interaccionan en absoluto. La presión de dicho gas es directamente proporcional a la temperatura e inversamente proporcional al volumen; la constante de proporcionalidad es R, la constante de los gases ideales. Esta relación se expresa matemáticamente por la ecuación $P = RT/V$, conocida como ley de los gases ideales.

Los métodos de perturbaciones sugieren una forma de modificar la ley de los gases ideales para describir las rela-

ciones existentes entre la presión, el volumen y la temperatura de los gases reales. A la ley anterior se añaden una serie de términos en los que la densidad del gas aparece elevada a potencias cada vez mayores. La teoría de perturbaciones de los gases imperfectos fue desarrollada con todo detalle por Joseph E. Mayer. Esta teoría da cuenta satisfactoria de los efectos experimentados siempre que la densidad no se haga demasiado grande. Cuando la densidad es elevada, los términos sucesivos de la serie no se hacen menores, sino mayores, y, en consecuencia, el cálculo resulta imposible. Por esta razón, el gas ideal no es un sistema de referencia satisfactorio para elaborar una teoría de los líquidos. La explicación de este fracaso radica en el hecho de que la estructura molecular de los gases ideales es muy diferente de la de los líquidos.

El modelo de esferas macizas, que constituye un sistema de referencia mucho más prometedor para los líquidos, da cuenta aproximada de las interacciones repulsivas que son importantes cuando la densidad es elevada, y, en él, las disposiciones más frecuentes de las moléculas que resultan son similares a las que hallamos en los líquidos.

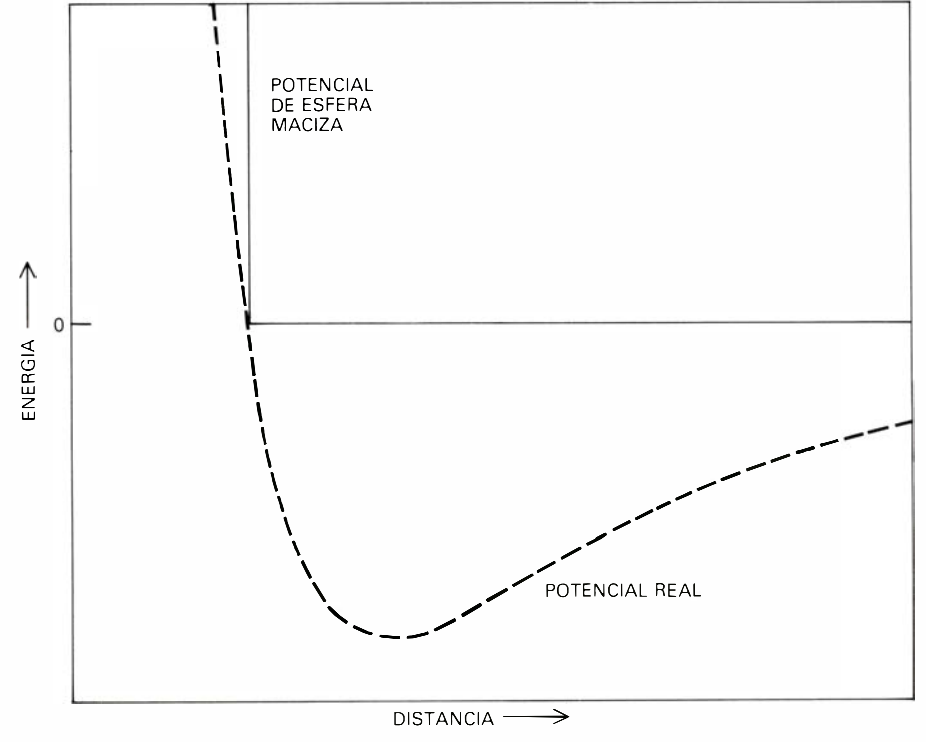
El primer intento de describir los líquidos utilizando el modelo de esferas macizas se debe a van der Waals, quien propuso que la ley de los gases ideales se modificara en dos aspectos. En primer lugar, debería reemplazarse el volumen V por la diferencia $V-b$, donde b representa el volumen del que una molécula resulta excluida en virtud de las interacciones repulsivas con otras moléculas. Partiendo de la suposición de que las moléculas son esferas macizas, se puede calcular el valor de la magnitud b. La segunda de las modificaciones propuestas consiste en sustraer del valor de la presión una cierta "presión interna", definida por la expresión a/V^2, donde a es una constante. El efecto de las fuentes atractivas se introduce a través de esta segunda modificación. En consecuencia, la ecuación que nos da la presión de un líquido responde a la forma $P = RT/(V-b) - a/V^2$.

Van der Waals justificó la sustracción de la presión interna en base a sus ideas intuitivas sobre la "asociación" de las moléculas. Van der Waals sugirió que, cuando una molécula se encuentra situada junto a las paredes de un recinto, tiene más moléculas detrás de ella que ante sí. Como consecuencia de ese hecho, las fuerzas atractivas de todas estas moléculas no se compensan entre sí y aparece una fuerza resultante que tiende a llevar a las moléculas consideradas hacia el interior del líquido, con lo cual la fuerza que esta molécula ejerce sobre la pared del recinto resulta ser menor. Una reducción en la fuerza que actúa sobre una pared equivale a una disminución de la presión. El efecto adquiere importancia tan sólo en el caso de los líquidos y gases comprimidos, es decir, en aquellos sistemas en los que la distancia mutua entre las moléculas es lo suficientemente pequeña como para que las fuerzas atractivas entre ellas sean significativas. Como resultado de todos estos hechos, la presión interna depende de la densidad.

Sobre la base de estas formulaciones, van der Waals demostró que se produce una transición continua entre los estados líquido y gaseoso. Sus ideas generales sobre la continuidad de los estados y el empaquetamiento de moléculas eran muy acertadas; sus propias aproximaciones específicas, tales como la expresión a/V^2 para la presión interna, no se alejaban mucho de la verdad. Sin embargo, van der Waals no dispuso de una teoría satisfactoria para los fluidos basada en el modelo de esferas macizas. Esta teoría no se desarrolló plenamente hasta hace unos 20 años, cuando se pudieron realizar los cálculos pertinentes con la ayuda de ordenadores.

El punto débil de la teoría de los fluidos de van der Waals radicaba en la magnitud b, representativa del volumen de exclusión. Si se supone que b es constante, la ecuación de van der Waals proporciona sólo un valor aproximado de la presión sobre la base del modelo de esferas macizas. Si la densidad es alta, la ecuación sobreestima el valor de esta magnitud debido a que los volúmenes de exclusión de las diferentes moléculas se solapan entre sí, de forma tal que el volumen de exclusión total es menor que la suma de los volúmenes excluidos por cada molécula considerada individualmente. En virtud de este hecho, los intentos de ajustar la ecuación de van der Waals a los datos experimentales suelen conducir a valores de b carentes de significado físico. El fracaso en la predicción de la presión oscureció el hecho de que van der Waals anduviera en lo cierto.

En las décadas de 1950 y 1960 varios investigadores empezaron a estudiar la



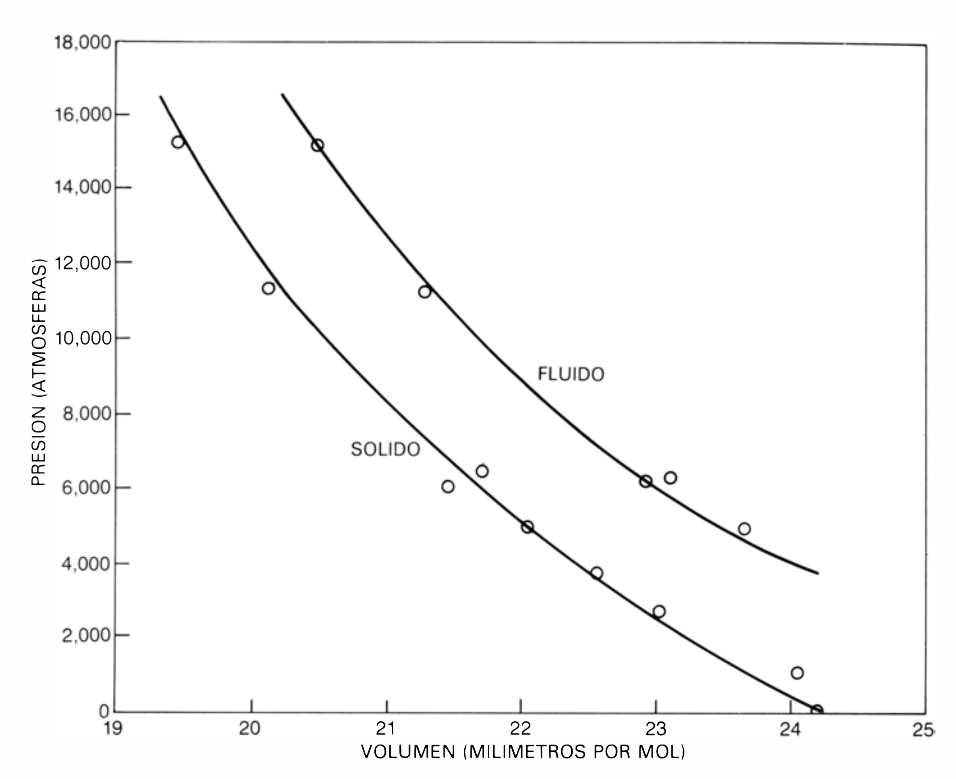
FUNCION DE ENERGIA POTENCIAL. A partir de ella se determina la fuerza que actúa entre dos moléculas; fuerza que se manifiesta como una tendencia de las moléculas a ajustar su separación a aquella distancia para la cual la energía potencial correspondiente se hace mínima. El potencial de un sistema de esferas macizas viene representado por una línea vertical para la distancia que corresponde al diámetro de las esferas y por una línea horizontal para todas las distancias mayores que la anterior. La línea vertical significa que la energía se hace infinita cuando se intenta comprimir dos moléculas para que su separación relativa sea menor que sus propios diámetros; en otras palabras, la existencia de una fuerza repulsiva infinita impide que las moléculas se solapen entre sí. La línea horizontal indica que entre las moléculas no actúa ninguna fuerza cuando su separación mutua es mayor que su diámetro. Una función de energía potencial real difiere del potencial de esferas macizas en dos aspectos: la parte repulsiva del potencial no es del todo vertical (se puede producir solapamiento entre las moléculas) y existe una región donde la energía potencial aumenta con la distancia; en esa región la fuerza intermolecular es atractiva.

aplicación del modelo de esferas macizas a la descripción de los fluidos reales. Hace quince años, nosotros mismos contribuimos a todo este trabajo actualizando las ideas originales de van der Waals en relación con el modelo de esferas macizas en el contexto de una teoría rigurosa y práctica al mismo tiempo. En nuestro trabajo utilizamos las ideas de Alder, E. B. Smith, J. S. Rowlinson, Donald A. McQuarrie y Joseph L. Katz. Por otro lado, John D. Weeks, David Chandler y Hans C. Andersen introdujeron posteriormente algunas mejoras en esta teoría.

En nuestra búsqueda de una formulación rigurosa de las ideas de van der Waals, procedimos a dividir el potencial intermolecular en dos partes de corto y largo alcance. La división se produce en el punto en que dicho potencial deja de ser atractivo y pasa a ser repulsivo. Las dos partes del potencial citadas anteriormente constituyen la base de los dos primeros términos de una serie de perturbaciones; la variable que se eleva a potencias cada vez más altas viene dada por la profundidad del potencial dividida por la temperatura. El primer término representa la abrupta repulsión de corto alcance, ligada al diámetro de las esferas macizas. El segundo término se refiere al efecto principal de la parte atractiva del potencial; la fuerza correspondiente es débil y de corto alcance.

Debido a que el potencial de corto alcance varía fuertemente con la distancia, se puede suponer que su contribución a la energía libre es casi la misma que la de la repulsión infinitamente abrupta que caracteriza a una esfera maciza ideal. Esta hipótesis es fundamental en el desarrollo de nuestra teoría. Con el fin de especificar el diámetro de dichas esferas de referencia, hemos desarrollado un criterio matemático que toma en consideración las interacciones repulsivas que se ponen de manifiesto en un líquido real.

La cuidadosa elección de un diámetro adecuado para las esferas macizas reviste una gran importancia debido al hecho de que las moléculas reales son más blandas que las esferas macizas que postula el modelo y sus funciones de energía potencial son algo "borrosas". De hecho, el diámetro de una molécula real no tiene un valor constante, sino que depende de la energía con que chocan las moléculas. En consecuencia, el diámetro efectivo de las moléculas viene dado por la media ponderada de la distribución de energía potencial repulsiva, de forma tal que las propiedades termodinámicas del sistema de esferas macizas correspondan a las propiedades de una sustancia real promediadas estadísticamente. El diámetro efectivo de las esferas macizas no es una constante y no depende de la densidad del sistema; por el contrario, dicho diámetro sí que depende de la temperatura o, lo que es lo mismo, de la energía media de las moléculas. La variación del citado diámetro efectivo permite que el potencial de las esferas macizas, muy simple pero un tanto irreal, se asemeje al potencial que gobierna las moléculas reales.



LAS FASES SOLIDA Y FLUIDA de una sustancia se pueden describir con precisión mediante un modelo en el que las interacciones entre las moléculas se definen a través de una función de energía potencial única. En esta figura se ha representado la relación existente entre la presión y el volumen para el argón; los puntos representan los datos experimentales y las líneas continuas constituyen los resultados de los cálculos teóricos realizados con la ayuda de un ordenador utilizando dicha función potencial.

El segundo término de nuestra ecuación de estado es una expresión de la energía total del potencial intermolecular promediado para todas las distribuciones moleculares posibles en el fluido de esferas macizas que nos sirve de referencia. Este término corresponde a la presión interna de la ecuación de estado de van der Waals y se puede determinar fácilmente a partir de la función de distribución radial que predice dicho modelo. El efecto fundamental de este término es disminuir la energía y la presión del líquido.

Ninguno de estos términos refleja el hecho de que las fuerzas atractivas existentes entre las moléculas, tal como fueron introducidas en las hipótesis de van der Waals, dan lugar a que la estructura de un líquido sea diferente de la que posee un fluido ideal de la misma densidad. Las fuerzas atractivas conducen a las moléculas hasta las regiones de energía más baja y, por consiguiente, reducen su energía libre. Para densidades elevadas, el efecto es pequeño dado que las moléculas se encuentran empaquetadas en forma compacta. Por el contrario, cuando la densidad es baja, el cambio en la entropía es apreciable. Un tercer término en nuestra ecuación da cuenta de los efectos de baja densidad y aparece como una pequeña diferencia en el valor de la entropía del fluido real y del fluido de referencia. Con la adición de estos tres términos nuestra suma de perturbaciones proporciona unos resultados satisfactorios para la energía libre de un fluido. El acuerdo entre el modelo teórico y la experiencia es excelente.

Cuando van der Waals aceptó el premio Nobel de física en 1910 hizo el siguiente comentario sobre la forma en que había introducido la presión interna en su teoría de los líquidos: "Confieso que todo esto es un rodeo. Tal vez exista un camino más directo. En la búsqueda de este camino, los *Principios básicos de Mecánica Estadística* de Gibbs serán una referencia obligada." La deducción de una teoría similar a la de van der Waals en los años 1960 y su constante clarificación y refinamiento, utilizando la teoría de perturbaciones y la mecánica estadística, se puede considerar como la culminación del trabajo que iniciara él hace más de un siglo.

Patrones de coloración de las mariposas diurnas y nocturnas

Los patrones alares son mosaicos de diminutas escamas. El despliegue de más de 100.000 patrones distintos está determinado por un número exiguo de normas básicas, que nos aproximan a la comprensión del desarrollo animal

H. Frederik Nijhout

Pocas cosas hay en la naturaleza que igualen la belleza y variedad de los patrones alares de las mariposas diurnas y nocturnas. El orden de los Lepidópteros (donde se encuadran esos insectos) abarca unas 100.000 especies; cada una de ellas distinguible del resto por el patrón de coloración de sus alas. Dato que resulta mucho más sorprendente cuando uno estudia (yo le dediqué varios años) y se interroga por la manera como se forman dichos patrones: la respuesta a la que se llega es de lo más simple.

Las bases fisiológicas de patrones tan diversos de coloración han venido preocupando desde hace tiempo a los biólogos, ya que representan un problema fundamental de la morfología animal. ¿Qué procesos regulan el desarrollo de los estereotipados patrones espaciales de tejidos, órganos y extremidades que caracterizan a determinada especie animal? ¿Qué diferencias esenciales en estos procesos causarán la apariencia característica de la especie? En principio, hay razones para creer que el desarrollo de los patrones de coloración pueda lograrse por procesos similares a los que guían el desarrollo de las características morfológicas, ya que todo desarrollo es, en último término, el resultado de cambios progresivos en la expresión de los genes. En los patrones de coloración, esta expresión da lugar a la síntesis localizada de moléculas de pigmento, que son relativamente sencillas.

Las escamas del ala de los Lepidópteros (lepidóptero significa alas escamosas) son bastante pequeñas y forman una capa sobre la membrana alar. Cada una es una segregación aplanada, de unas 100 micras de longitud por 50 de anchura, procedente de una sola célula. (Una micra es la milésima parte de un milímetro.) Las escamas se superponen como las tejas en un tejado, cubriendo completamente la superficie membranosa del ala. Su densidad varía entre 200 y 600 unidades por milímetro cuadrado. Cualquiera que haya aprehendido una mariposa por las alas sabe que las escamas se desprenden fácilmente, y que con ellas desaparece el patrón de coloración. Queda entonces al descubierto la naturaleza incolora y translúcida de la membrana alar.

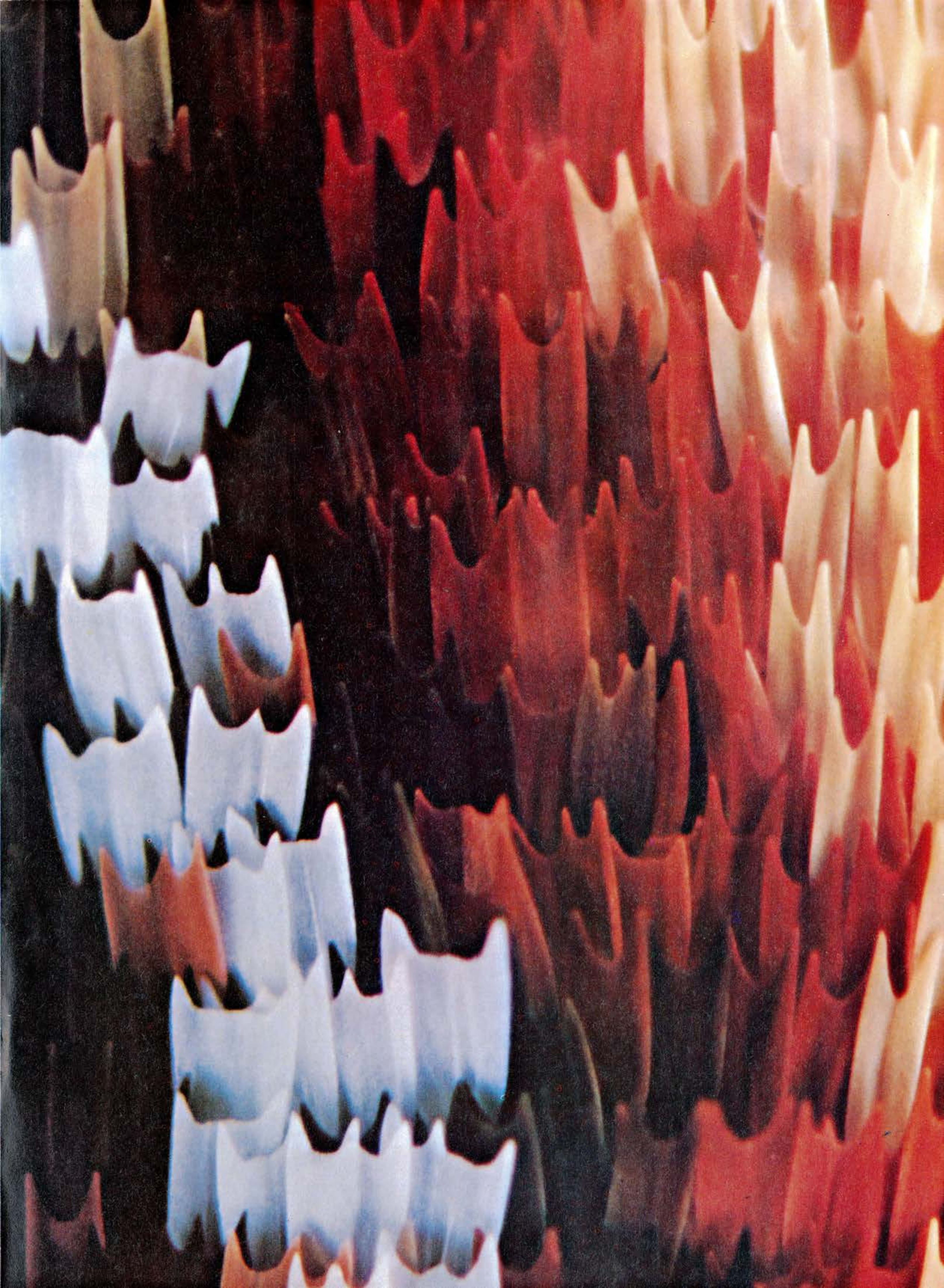
Al examinar un ala a gran aumento vemos claramente que el patrón de coloración es un perfecto mosaico de tejas. Cada teja del mosaico es una escama, y cada escama es de un solo color. Los patrones habituales comprenden pocos colores: tres, cuatro o cinco, por regla general. Las gradaciones del color y la tonalidad se consiguen variando las proporciones de las diferentes escamas, de igual modo que las fotografías de este artículo son el resultado de una combinación de puntitos de sólo cuatro colores. Así, el patrón del ala anterior de la mariposa *Precis coenia* resulta de la combinación de escamas de cuatro colores: negro, ante, rojo y marrón. Las escamas blancas del centro de la gran "mancha ocelar" son incoloras; su blancura se debe a la reflexión de la luz por parte de diminutas burbujas de aire contenidas en las escamas. Aunque cada escama tenga sólo un pigmento, su cantidad puede variar. En la mariposa *Precis coenia* se han detectado microscópicamente un mínimo de tres tonalidades en las escamas de color marrón. Igualmente, las escamas ante y rojas parecen tener el mismo pigmento, aunque está menos concentrado en las ante que en las rojas.

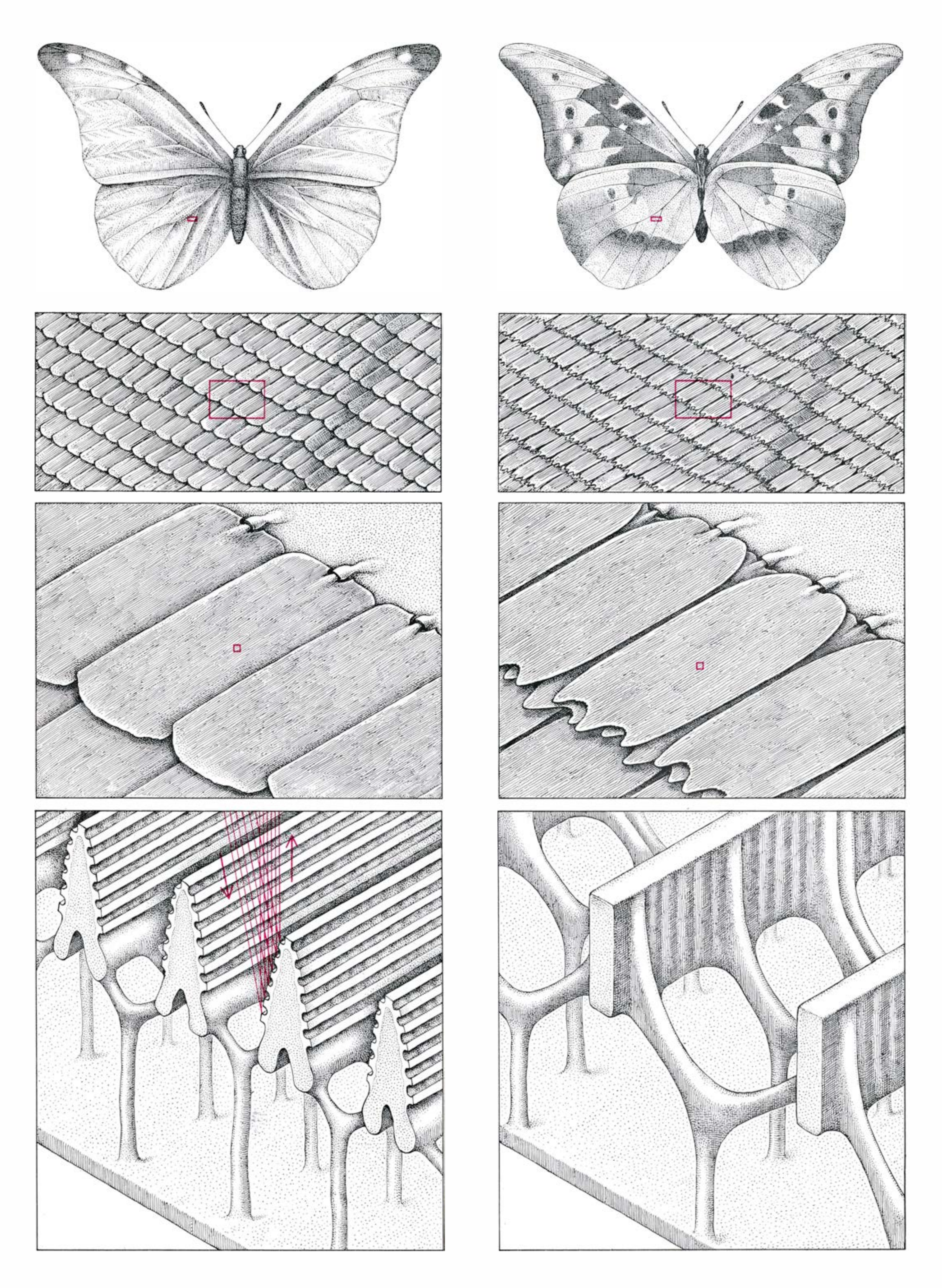
Los colores de las escamas de mariposas diurnas y nocturnas son de dos tipos: pigmentarios y estructurales. Los colores pigmentarios se deben sólo a unos pocos productos químicos; en general, melaninas. Así, todos los negros y grises y la mayoría de tostados, marrones, pardo-rojizos y amarillentos son formas de melanina. La segunda gran categoría de pigmentos alares son las pterinas, que dan lugar a la mayoría de rojos brillante, naranjas y amarillos y algunos de los blancos. Se han detectado en ciertas especies una serie de pigmentos no identificados, algunos de ellos caracterizados como flavonas, carotenoides y ommocromos; son responsables de la gama que va del amarillento al parduzco.

De los colores más espectaculares hay algunos que no tiene nada que ver con productos pigmentarios, sino que son fruto de peculiaridades estructurales de las escamas. Hemos aludido ya al blanco de los ocelos. Además, todos los colores iridiscentes o metálicos, así como los azules y verdes de las mariposas diurnas (excepto los verdes de las nocturnas), son estructurales. Estos colores resultan de la interferencia, difracción o dispersión de la luz.

Las escamas iridiscentes deben esencialmente su color a la interferencia de la luz reflejada. Según la especie, se forman colores de interferencia al reflejarse la luz en las crestas horizontales de las numerosas aristas verticales de la superficie de la escama, o en capas ho-

ESCAMAS ALARES de la mariposa nocturna *Hyalophora cecropia*, aumentadas 150 veces en la microfotografía de la página contigua. El área mostrada se encuentra en las proximidades de la "mancha ocelar", que puede apreciarse cerca del extremo de cada una de las dos alas anteriores del insecto. Las escamas blancas son en realidad incoloras, y deben ese tono a la luz reflejada por diminutas burbujas de aire que encierran en su interior. El color de las escamas, que va del pardo-rojizo al ante pasando por el rojo, obedece a la presencia de distintas cantidades del mismo pigmento ommocromo. Fotografía de F. Goro.





rizontales de delgadas lamelas que constituyen la cutícula externa de la escama. Las escamas iridiscentes poseen también un pigmento, generalmente melanina oscura, que absorbe la mayor parte de la luz transmitida, dando lugar a que el componente reflejado brille con especial intensidad.

Se diría, pues, que los diversos patrones de coloración de las alas de lepidópteros siguen un conjunto de reglas de diseño muy claras. Estas reglas fueron formuladas por vez primera en los años veinte, de manera independiente y casi simultánea, por B. N. Schwanwitsch, de la Universidad de Petrogrado (hoy Leningrado), y Fritz Süffert, de la Universidad de Friburgo. Observaron que la mayoría de los patrones de coloración podían derivarse de un solo arquetipo, o plano general, al que se denomina habitualmente plano ninfálido, ya que tanto Schwanwitsch como Süffert lo describieron por vez primera en los Nymphalidae, una numerosa familia de mariposas.

Con los años ha ido quedando patente que, a partir de este plano general, puede derivarse la mayoría de patrones alares de las principales familias de mariposas diurnas y nocturnas. Sin embargo, es preciso señalar que el plano general no ha de contemplarse como patrón de coloración ancestral, sino como un patrón hipotético a partir del cual podemos obtener, mediante alteraciones apropiadas, los patrones de especies existentes.

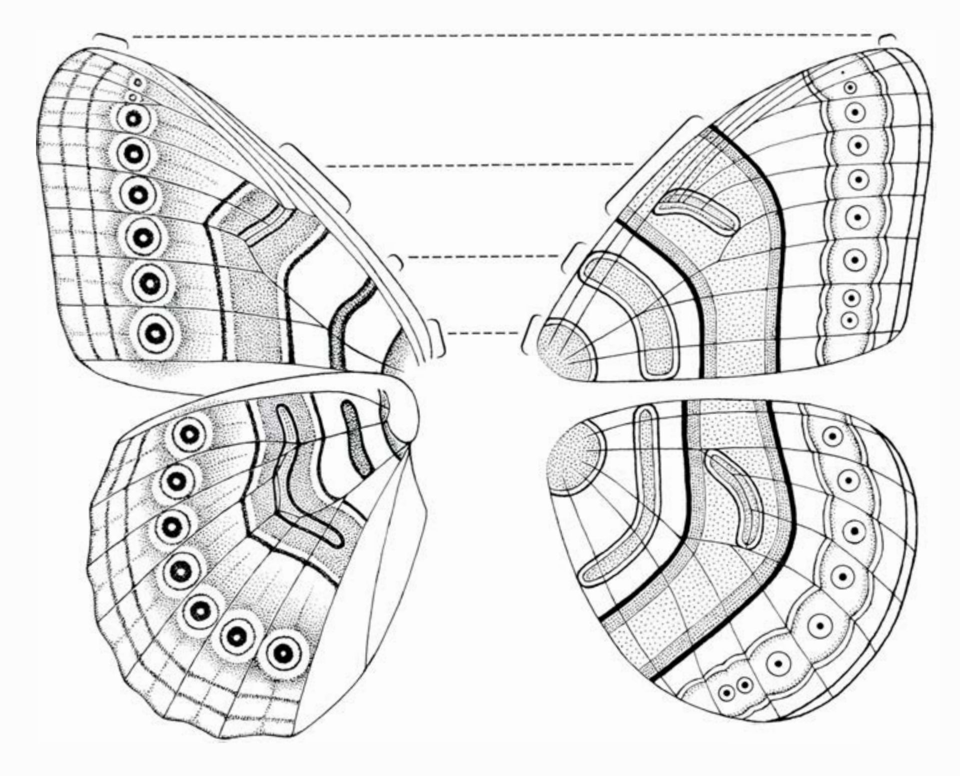
Las características principales del plano general ninfálido son las siguientes. A lo largo del margen del ala existe una hilera de ocelos, o pequeñas manchas ocelares, formados por una serie de círculos concéntricos con sus centros siempre situados en la línea media de una celda alar. (La expresión celda alar designa el área de la membrana alar limitada por venas.) El plano general contempla un ocelo en cada una de las celdas alares periféricas.

En la región media del ala hay bandas coloreadas que van del margen anterior al posterior del ala. Constituyen el llamado sistema de simetría central, que consta de dos conjuntos de bandas pigmentarias situadas a izquierda y derecha de una línea imaginaria que discurriría por el medio del ala. La secuencia de coloración en un lado es imagen especular de la existente en el lado opuesto; consiguientemente, los colores de las bandas son simétricos respecto a la línea media del ala.

Una franja notablemente pigmentada, la mancha discal, se observa por lo general en la línea media del sistema de simetría central. Su localización suele definir el eje de simetría y en muchas especies de mariposas nocturnas aparece como una mancha ocelar grande y rendondeada. Finalmente, podemos encontrar cerca de la base del ala uno o más sistemas de simetría basal formados por conjuntos armónicos de bandas coloreadas, similares a las del sistema de simetría central, aunque bastante menos elaboradas que en éste.

Ninguna especie conocida presenta el plano general completo. Por plano general se entiende el patrón extremo a partir del cual podemos derivar los patrones existentes por expresión o exageración de algunos elementos y supresión de otros. Así, los ocelos marginales se presentan en realidad en unas cuantas celdas alares y la banda proximal del sistema de simetría central (la más próxima al cuerpo) falta a menudo. Por si fuera poco las bandas del sistema de simetría central pueden tornarse harto irregulares, y no parece existir relación entre las formas de los elementos distales (los más apartados del cuerpo) y los proximales. Todo elemento del patrón puede expresarse en una amplia gama de pigmentos o colores estructurales, dando lugar a grandes contrastes o a sutiles gradaciones.

La modificación principal del plano general implica desplazamientos laterales de regiones concretas de las bandas del sistema de simetría central. Schwanwitsch llamó dislocación al fenómeno porque su efecto visual le recordaba la dislocación geológica, donde las capas de rocas sedimentarias se han desplazado a lo largo de una fractura. En los Lepidópteros, la dislocación implica la interrupción de una banda pigmentada cuando cruza una vena alar y el desplazamiento lateral de una parte importante de la banda a lo largo de la



PLANO GENERAL NINFALIDO, esquema propuesto independientemente por B. N. Schwanwitsch, de la Universidad de Petrogrado (Leningrado), y Fritz Süffert, de la Universidad de Friburgo, en los años veinte, del que pueden derivarse la mayoría de patrones alares de mariposas diurnas y nocturnas. Los planos aquí mostrados, el de Schwanwitsch a la izquierda y el de Süffert a la derecha, se esquematizan en sus semejanzas y diferencias. Cerca del 75 por ciento de los patrones de coloración de las mariposas pueden derivarse de estos planos generales por exageración de algunos elementos y supresión de otros.

ESTRUCTURA DETALLADA de las escamas de la mariposa sudamericana *Morpho rhetenor*, desentrañada en la serie de dibujos de aumento progresivo que aparece en la página contigua. A la izquierda, los dibujos muestran la cara superior del ala, que en el macho es azul metálico. Se trata de un color estructural, ya que surge de la interacción de la luz con las escamas alares. Las escamas azules deben su color a los detalles estructurales de las aristas de su superficie. Estas aristas presentan crestas horizontales cuyas superficies superior e inferior distan unos 0,22 micrometros, es decir, aproximadamente la mitad de la longitud de onda de la luz azul, lo que da lugar a que las longitudes de onda azules de la luz que se refleja en la superficie interfieran positivamente. A la derecha, los dibujos representan la cara inferior del ala de la mariposa, que es de color marrón. Este color resulta de los pigmentos del esqueleto de la escama.

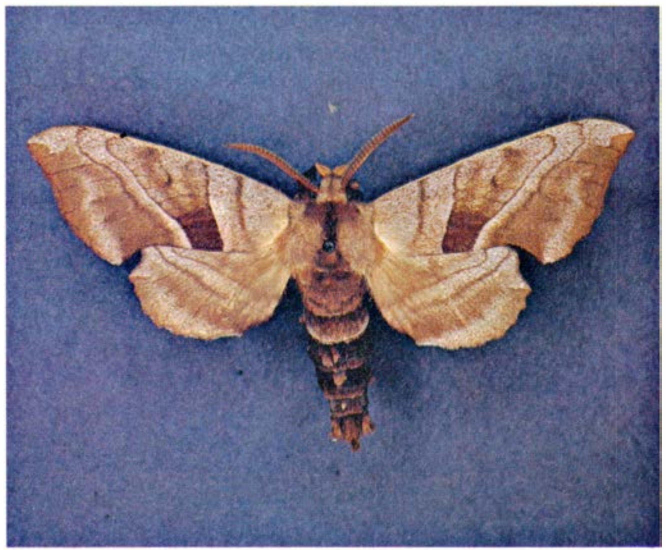
vena. Patrones de dislocación característicos de cada especie dan cuenta de la mayoría de las variaciones de los patrones alares. En numerosos casos, el efecto dislocador genera disposiciones discontinuas de rayas y puntos coloreados. También se habla de dislocaciones cuando los ocelos no guardan perfecta alineación.

El estado actual de nuestros conocimientos nos permite afirmar que muchos aspectos del diseño y desarrollo del patrón de coloración de los Lepidópteros hallan su explicación en el marco conceptual de la biología del desarrollo. A lo largo de la última década muchos biólogos del desarrollo han aceptado la idea de Lewis Wolpert, del Middlesex Hospital Medical School de Londres, de que las células deben tener información sobre su posición para que el desarrollo proceda normalmente [véase "La formación de modelos en el desarrollo biológico" por Lewis Wolpert; Investigación y Ciencia, diciembre de 1978]. En consecuencia, debe existir un mecanismo mediante el cual las células "conozcan" su posición en un tejido u órgano en desarrollo, sirviéndose de esa información posicional para activar o inhibir luego determinados procesos bioquímicos especificados por sus genes y que corresponden a las células de esa posición. En el desarrollo de los patrones de coloración ello significa que una célula escamosa debe saber cuál es su posición en el ala y producir el pigmento que corresponde a tal ubicación.

Al formular su teoría de la información posicional, Wolpert sugirió que puesto que las células sólo podían saber dónde se hallaban refiriéndose a algo externo a ellas, todo sistema en desarrollo debía tener como mínimo un punto de referencia. Wolpert no especificó la naturaleza bioquímica o fisiológica de esos puntos; de ahí que buena parte de las investigaciones teóricas se hayan dirigido a definir qué tipos de procesos o fenómenos pudieran ejercer esa misión. Casi todos los modelos propuestos presuponen la existencia de algún tipo de gradiente: un factor físico o una sustancia química que cambian gradualmente en su dimensión o concentración dentro de un campo de desarrollo. Los modelos que admiten una representación más sencilla proponen la









MODIFICACIONES DEL PLANO GENERAL. La mariposa diurna *Stichophthalma camadeva* (*arriba a la izquierda*) es la especie con elementos más similares a los del plano general. El gran satúrnido *Saturnia carpini* (*arriba, a la derecha*) posee un sistema de simetría central patente y una mancha discal bien desarrollada, pero sólo un esbozo de ocelos ("manchas ocelares") marginales. El esfíngido *Cressonia juglandis* (*abajo, a la izquierda*) presenta, al igual que muchas mariposas nocturnas, un patrón menos manifiesto, aunque destaca claramente el sistema de simetría central. La mariposa tropical *Taenaris rothschildia* (*abajo, a la derecha*) presenta un patrón de coloración reducido, con los ocelos marginales desarrollados de una forma desproporcionada.

existencia de fuentes discretas de una sustancia química difusible: un morfógeno. La concentración del morfógeno sería máxima en las células próximas a la fuente, para decrecer gradualmente en las células cada vez más alejadas de ella. En razón de su concentración, el morfógeno afectaría diferencialmente la expresión de determinados genes de las células por donde se difunde.

Aceptemos de momento esta idea. Supongamos que los genes implicados codifican enzimas de las vías biosintéticas de varios pigmentos. En estas circunstancias, los patrones más simples que podrían formarse serían anillos y círculos. Lo único que se exige es que las células que se hallen dentro de la fuente de información posicional, o a determinada distancia crítica, se diferencien de un modo idéntico. Sólo así todas las células sintetizarían el mismo pigmento.

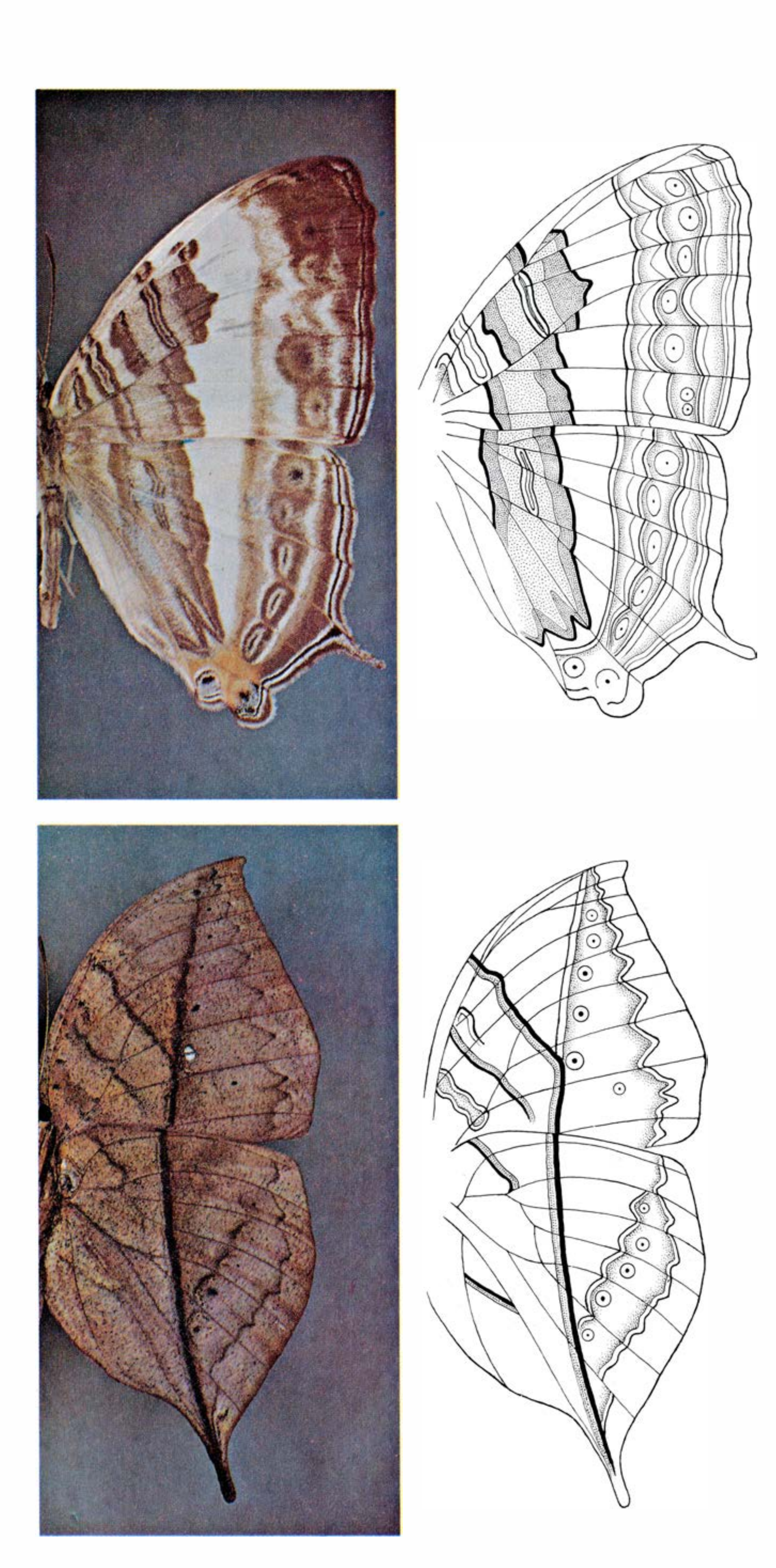
Hemos señalado ya que los círculos y anillos de pigmento constituían rasgos comunes de los patrones de coloración de los Lepidópteros, manifestándose como ocelos marginales y manchas ocelares discales. Podríamos ahora aprovechar la idea de información posicional y utilizarla de modelo para analizar el desarrollo de las manchas ocelares. Según este modelo, el desarrollo de una mancha ocelar circular dependería, en cierto modo, de la actividad especializada de las células centrales. (Interesa señalar que casi todas las manchas ocelares presentan un área central muy nítida, de coloración más clara que los círculos de pigmento que la rodean.)

No se pone ya en cuestión que los procesos fisiológicos que determinan la localización y forma del patrón de coloración acontezcan cuando el ala no esté aún pigmentada. De ahí que uno de los problemas prácticos más arduos que se presentan al experimentar sobre la determinación del patrón de coloración sea predecir con exactitud en qué lugar del ala aparecerá un elemento del patrón. He dedicado varios años a estudiar el desarrollo del patrón de coloración de la mariposa *Precis coenia*, especie particularmente útil para análisis de ese tipo, ya que presenta varios marcadores en la cubierta alar de su pupa, un estadio del desarrollo anterior a la aparición del adulto. Con la ayuda de tales marcadores podemos predecir de un modo exacto la posición de varias partes del presunto patrón de coloración en el ala no pigmentada mucho antes de que aparezca el patrón real.

Resulta fácil identificar la posición que hipotéticamente ocupará el área central de la gran mancha ocelar del ala interior. Sirviéndonos de una aguja de cauterización atravesaremos la cutícula pupal y eliminaremos el pequeño grupo de células del centro de la mancha ocelar. Si se realiza la operación pocas horas después de la pupación, en el instante inmediato tras la formación del ala, se inhibe total, y específicamente, el desarrollo de la mancha ocelar. Cauterizaciones similares realizadas a las 48 o más horas de la pupación no tienen efecto alguno, mientras que si eliminamos las células en distintos períodos dentro de las primeras 48 horas después de la pupación se obtienen manchas ocelares de tamaños intermedios. La pigmentación del ala no se manifiesta hasta el quinto día del estado de pupa (aproximadamente unas 24 horas antes de la aparición del adulto); por tanto, la determinación del patrón precede en unos tres días a la diferenciación del patrón (síntesis de pigmento).

Si el desarrollo de la mancha ocelar depende de alguna propiedad especial de las células del centro de la mancha, será posible, trasplantando las células, transferir esa propiedad a otros lugares del ala. La transferencia da resultado. Al trasplantar un trocito de epidermis del área que hubiera formado el centro de una mancha ocular a una región distinta del ala (de la que se ha extirpado un trozo de epidermis de similar tamaño) encontré que se desarrollaba, a su alrededor, un patrón pigmentario circular. Los trasplantes de epidermis de otras regiones del ala no reprodujeron el patrón. Parece, pues, que la determinación de la mancha ocelar constituye una propiedad intrínseca de un pequeño número de células (quizá sólo una), que representan una de las fuentes de información posicional postuladas por Wolpert, al menos para uno de los elementos del patrón de coloración. A ese grupo de células lo llamaré foco en adelante.

Hemos señalado que el ala anterior de la mariposa *P. coenia* tenía sólo cuatro pigmentos, melaninas todos ellos. Aunque se desconocen las diferencias moleculares existentes entre estas distintas sustancias pigmentarias, hay pruebas que indican que, para su síntesis a partir de un precursor simple, se precisa un solo enzima: una tirosinasa. Parece también que para cada uno de los cuatro colores de la melanina se emplearían tirosinasas con distintas especificidades. De ser ello cierto, la acción del foco resultaría en principio muy simple: provocar, de alguna manera,



DESVIACIONES DEL PLANO GENERAL, que reflejan la flexibilidad del sistema. Generalmente, la variabilidad se debe a dislocaciones, término con el que Schwanwitsh designó a la interrupción de una banda de pigmento al cruzar una vena alar y desplazamiento lateral a lo largo de la vena. Mostramos aquí, en la fotografía del ala y el dibujo adjunto, dos de estos ejemplos. En la parte superior, la mariposa *Cyrestis cocles* muestra dislocaciones menores, aunque evidentes, de las bandas del sistema de simetría central. En la parte inferior, la mariposa-hoja de la India *Kallima inachus* ha desarrollado un patrón que semeja una hoja seca en el lecho de la jungla, patrón que incluye manchas mohosas y venas alares.

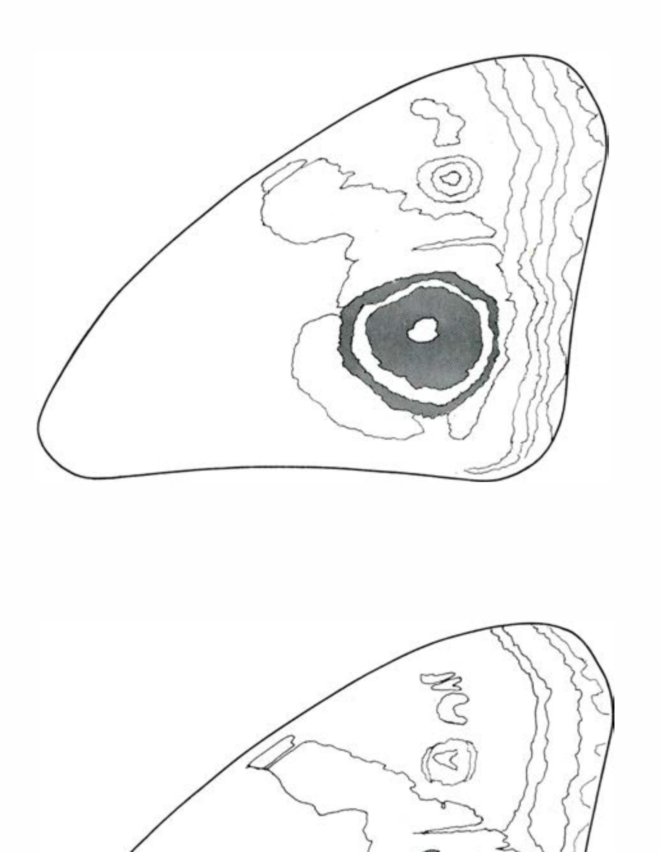
que las células escamosas bajo su dominio alteraran, por ejemplo, la producción de una tirosinasa de melanina amarilla en favor de una tirosinasa de melanina negra.

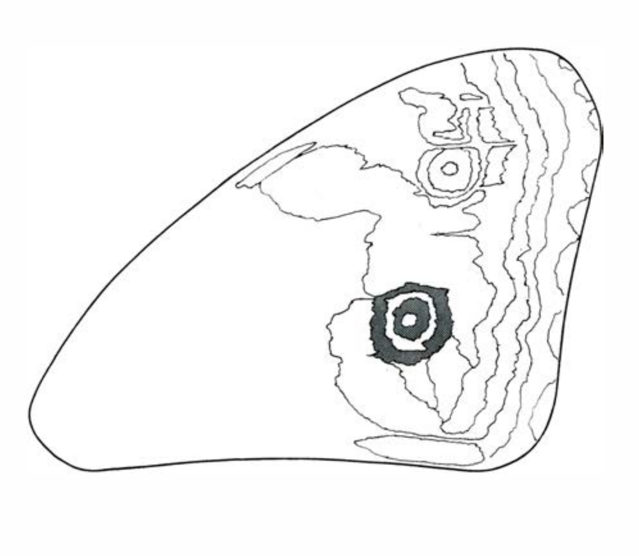
Es probable que todas las manchas ocelares se organicen alrededor de focos de este tipo y que el sistema de ocelos marginales surja alrededor de una hilera de focos. En el plano general ninfálido habría un foco por celda alar, situado en la línea media de la celda.

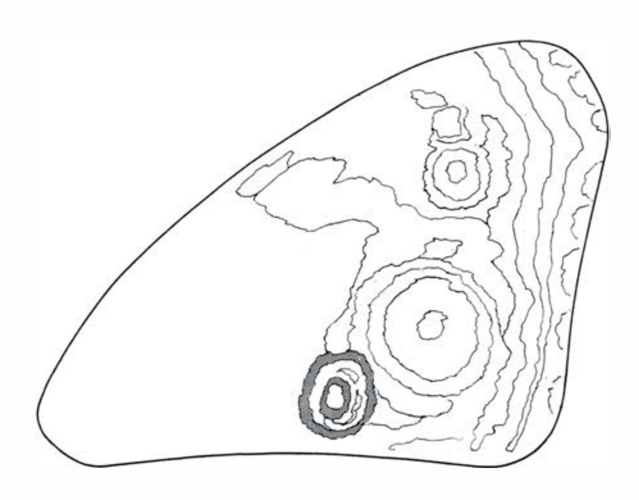
Hay buenas razones para creer que el sistema de simetría central se forma también alrededor de una hilera de focos. En muchas especies, este sistema consiste en una serie de círculos concéntricos fusionados y, al igual que en los ocelos marginales, los centros de los círculos se hallan en la línea media de las celdas alares. Es más, pruebas experimentales muestran la existencia de focos funcionales en el sistema de simetría central. En la década de 1930, Alfred Kühn y Melitta von Engelhardt, de la Universidad de Göttingen, realizaron una serie de experimentos en los que estudiaban el efecto de la cauterización en diversos puntos del sistema de simetría central de la polilla mediterránea de la harina (*Ephestia kühniella*). Analizando los defectos aparecidos en el patrón de coloración después de las cauterizaciones, revelaron que el sistema de simetría central de esta polilla respondía a la existencia de dos, o probablemente tres, "fuentes de determinación": una cercana al margen anterior del ala, otra próxima al margen posterior y otra quizás en el centro del ala. Experimentos similares realizados por otros investigadores en especies distintas confirmaron aquellos resultados.

Al igual que el sistema de simetría central, los sistemas de simetría basal se descomponen en pequeñas series de círculos concéntricos fusionados, cada uno con su centro situado en la línea media de una celda alar. El plano general ninfálido parece así organizado en torno a tres hileras de focos que discurren del margen anterior al posterior del ala. Una hilera distal de focos controla el desarrollo de los ocelos marginales, otra, central, el sistema de simetría central y una tercera, proximal, controla el sistema de simetría basal. El sistema de simetría central y los ocelos marginales constituyen, pues, distintas manifestaciones de unos procesos similares en lo fundamental.

Podría pensarse que las bandas de un sistema de simetría se originan por fusión de los anillos periféricos de gran-







DOS EXPERIMENTOS muestran el control del desarrollo de una mancha ocelar en una mariposa diurna o nocturna. Los experimentos corresponden a la mariposa *Precis coenia*. El patrón normal de la mariposa (*arriba, a la izquierda*) se altera profundamente al cauterizar, poco después de la formación de la pupa (estadio del desarrollo previo a la aparición del adulto), un pequeño grupo de células del centro de la mancha ocelar. Si la cauterización se realiza dentro de las cuatro horas que siguen a la pupación, la mancha se desarrolla muy incompletamente (*arriba, a la derecha*). Cauterizaciones realizadas 24 horas después dan lugar a una mancha ocelar de mediano tamaño (*abajo, a la izquierda*), mientras que cauterizaciones realizadas pasadas ya las 48 horas no producen efecto alguno. Si en vez de eliminar las células centrales, las trasplantamos a otra área del ala, aparecerá un patrón de círculos (*abajo, a la derecha*). La estructura formada no es una mancha ocelar completa, ya que el tiempo necesario para que el tejido trasplantado cicatrice adecuadamente determina que el período de inducción necesario para el desarrollo a término de una mancha ocelar sea inferior a las 48 horas prescritas.

des ocelos. Contradice esta idea el hecho de que las bandas del sistema de simetría central suelen ser líneas rectas, y no segmentos de círculos. Frecuentemente los ocelos marginales presentan, además, una nítida configuración no circular.

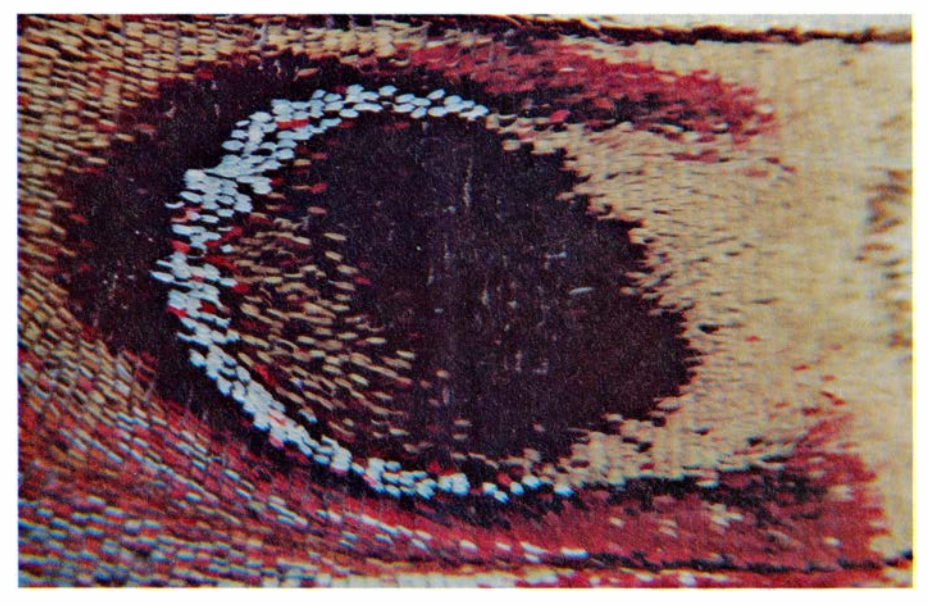
Antes de enfocar el problema del desarrollo de patrones no circulares alrededor de un foco conviene resumir las reglas que, en mi opinión, controlan el desarrollo del patrón alar en los Lepidópteros. Por lo que yo sé, los patrones de las superficies superior e inferior del ala se desarrollan de una forma independiente. Los elementos del patrón se desarrollan alrededor de un número finito de focos. En la mayoría de los casos, el foco determina tan sólo el patrón de la celda alar donde se halla. De esta manera, el patrón de coloración de cada celda alar se determina independientemente de los patrones de las demás celdas alares. Así pues, además de constituir un mosaico de escamas, el patrón de coloración integra un mosaico de celdas alares, que son unidades independientes de desarrollo. Cabe que un foco determine el patrón de varias celdas alares contiguas, pero sólo si éstas carecen de focos activos.

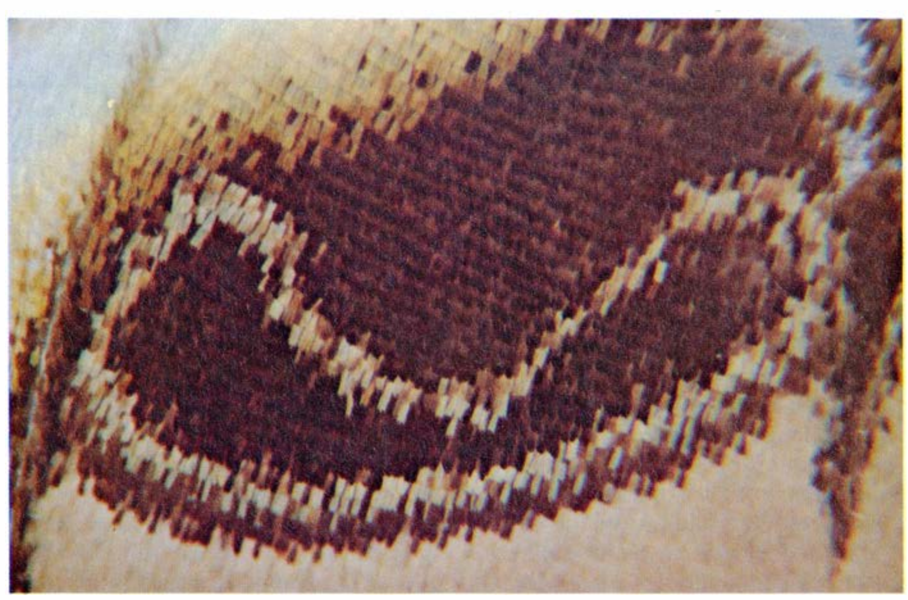
Dentro del plano general ninfálido, los patrones propios de cada especie se originarían, al menos en parte, variando el número y posición de los focos que se expresan. Así, el fenómeno de dislocación podría explicarse por desplazamientos laterales de los focos, que alterarían el patrón propio de cada celda alar. Sin embargo, los patrones propios de cada especie también dependerían de las variaciones de las reglas que controlan la forma del patrón que surge alrededor de un foco determinado.

Aunque las alas de los Lepidópteros presentan numerosos círculos, éstos se hallan en minoría con respecto al resto de elementos de los patrones alares. A menudo, los ocelos marginales se prolongan en elipses, se doblan en crecientes o se modifican en formas discontinuas más complejas. Parecidas formas no circulares se observan en las bandas del sistema de simetría central. Un aspecto interesante de los patrones no circulares es que casi siempre presentan simetría bilateral respecto al eje de simetría que, paralelamente a las venas alares y a través de un foco, discurre por la línea media de la celda alar. Generalmente, los patrones no circulares, al igual que los ocelos redondeados, quedan restringidos a una sola celda alar.

Como se desprende de la propia noción de información posicional, son dos las condiciones necesarias para el desarrollo de patrones circulares. En primer







PATRONES NO CIRCULARES que surgen con frecuencia alrededor de un foco (fuente de información posicional para aquellas células que producen pigmento). De arriba abajo se muestran los casos de las mariposas *Euphydryas editha, Hyalophora cecropia* y *Rothschildia arethusa*. El desarrollo de patrones de este tipo indica que la señal del foco no es interpretada por igual por todas las células que lo rodean.

lugar, la señal posicional ha de transmitirse desde el foco a igual velocidad e intensidad en todas direcciones. En segundo lugar, todas las células del ala deben seguir reglas similares al interpretar la intensidad de la señal posicional. Cuando se forma un patrón no circular alrededor de un foco resulta inmediato que no se cumple alguna de estas condiciones.

En su versión más estricta, la noción de información posicional de Wolpert supone que la señal posicional es invariable. Numerosas observaciones y experimentos apoyan esta aserción. En este caso, existiría una señal posicional única, o unas pocas clases de señales universales, y las diferencias en el estado genético y en los pasos previos del desarrollo de las células que responden determinarían el patrón que se desarrolla. En ese contexto, el desarrollo de patrones no circulares requeriría que las células escamosas de lugares distintos del ala siguieran diferentes reglas de interpretación de la señal posicional procedente del foco.

Pueden imaginarse numerosos mecanismos que den lugar a la formación de patrones de simetría bilateral alrededor de un foco. El modelo que yo sustento, que tiene a su favor una fácil visualización y, más importante todavía, ser verificable, implica una interacción de dos gradientes.

Para entender el modelo es necesario figurarse una imagen más definida de la señal de información posicional que emana del foco. Imaginémosla como una sustancia química, *P*, sintetizada en el foco y difundiéndose por igual en todas direcciones. El gradiente tridimensional de concentración de *P* sería un cono, con una concentración máxima en el foco que descendería gradualmente al alejarnos de él. La forma exacta del gradiente dependería de diversos factores, aunque, al no afectar al modelo, podemos suponerla lineal.

Si las células escamosas están programadas para sintetizar un pigmento negruzco dondequiera que *P* alcance una determinada concentración, el resultado será la aparición de un anillo de pigmento negruzco. Podemos visualizar esta situación cortando el gradiente cónico de *P* con un plano horizontal y proyectando la sección cónica obtenida sobre la superficie del ala. La altura del plano corresponderá a la concentración que determine la síntesis del pigmento negruzco. Podemos representar pigmentos distintos mediante planos diferentes, siendo la anchura del anillo de pigmento proporcional al grosor del plano correspondiente. De esta mane-

ra, los planos se tornan representaciones gráficas de las reglas mediante las cuales se interpreta el gradiente de información posicional.

Basándonos en el mismo gradiente de información posicional, podemos explicar los patrones que se desvían de la circularidad si damos por supuesto que las células escamosas situadas a un lado del foco están programadas para sintetizar un determinado pigmento a unas concentraciones de P superiores a las que lo hacen las células escamosas del lado opuesto. Si, en una representación tridimensional, un plano horizontal indica que el gradiente de P se interpreta por igual en cualquier lugar del ala, una regla útil será considerar que la altura de la superficie transecta es proporcional a la concentración P para la que se sintetiza un determinado pigmento. De esta manera, se formaría una elipse cuando el plano fuera inclinado.

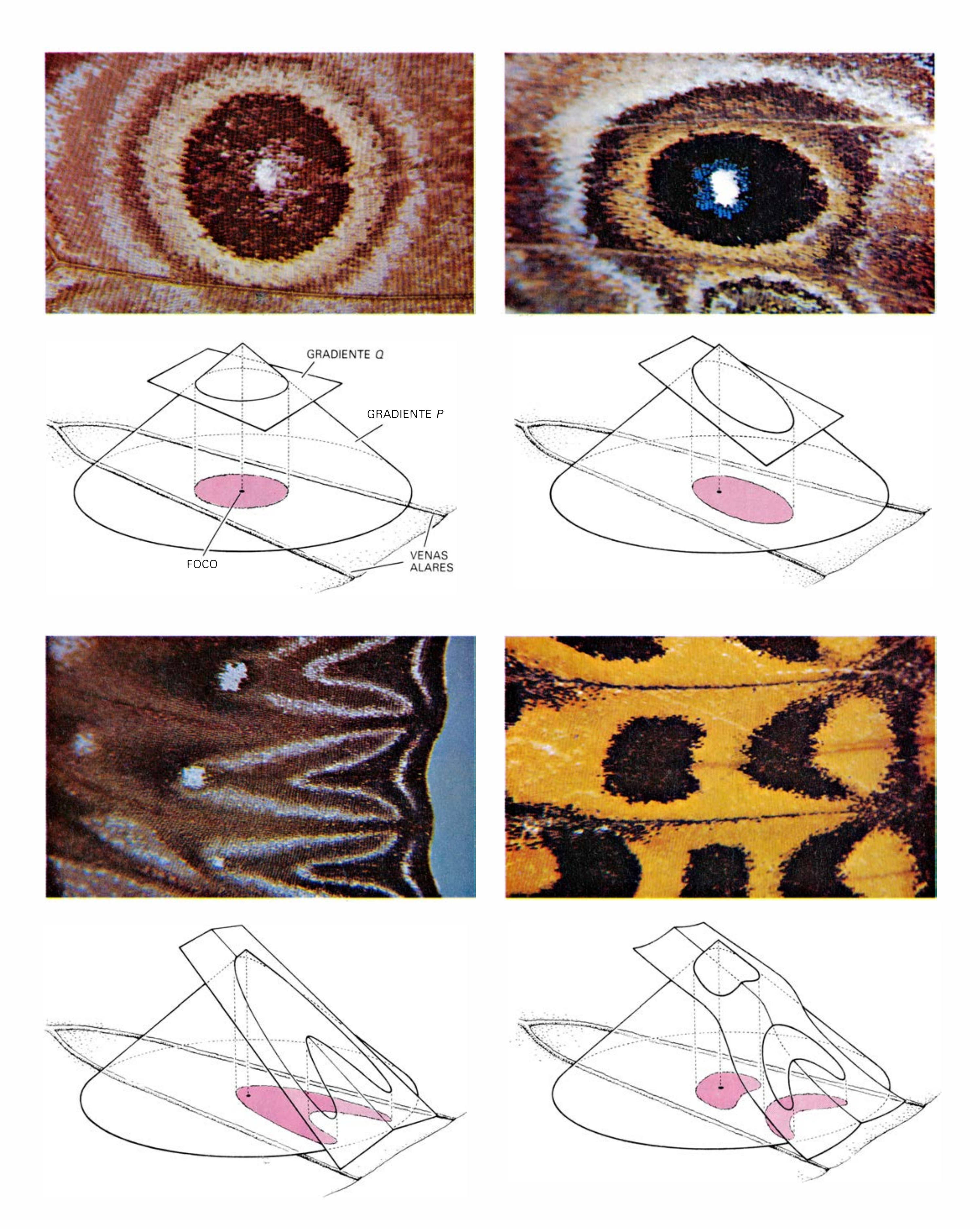
Llegados a este punto, se siente uno instado a pensar que el plano transversal sea un gradiente de otra sustancia química, Q, producida cerca de la base de la celda alar y que se difunde distalmente. Si suponemos entonces que el pigmento se sintetiza sólo cuando existe una determinada proporción entre las concentraciones de los dos productos, no será necesario disponer de un producto distinto para especificar la posición de cada pigmento en un patrón, ya que una proporción distinta entre los dos productos daría lugar al mismo resultado.

Desgraciadamente, apenas sabemos nada de la fisiología del desarrollo de patrones no circulares; tan sólo que los focos intervienen de algún modo. Además, la existencia de un segundo gradiente pertenece todavía al terreno de la hipótesis, pero ésta señala ya un camino para la investigación futura. De haber tal gradiente, podríamos alterarlo por extirpación quirúrgica de su fuente. Utilizando focos trasplantados podría experimentarse también en distintas regiones de este segundo campo-gradiente y observar si los resultados concuerdan con las predicciones formuladas sobre las formas de los patrones que aparecían alrededor de un foco trasplantado "hacia arriba" o "hacia abajo" del gradiente.

En muchas especies de Lepidópteros los ocelos marginales han perdido toda similitud con elipses o círculos, apareciendo en forma de flechas, galones, o crecientes, con su lado convexo dirigido a menudo hacia el foco y siempre con simetría bilateral respecto a la línea



EL PATRON GENERAL del ala de una mariposa diurna o nocturna es un mosaico de patrones independientes, uno por cada celda alar (una celda alar es un área limitada por venas). El patrón de la mariposa *Smyrna blomfildia* se muestra entero en la parte superior y descompuesto en celdas alares en la parte inferior. Los patrones de las caras superior e inferior del ala se desarrollan también independientemente.



MODELO DE DOS GRADIENTES propuesto como mecanismo generador de los ocelos marginales. El modelo está ilustrado por fotografías de los patrones alares de cuatro especies y sus correspondientes diagramas, debajo, que muestran a los dos gradientes, *P* y *Q*, en forma de un cono y un plano, respectivamente, que se intersecan. La forma tridimensional de un gradiente lineal (formado por un producto que difunde por igual en todas las direcciones de la celda alar) sería un cono, con la fuente del producto, en este caso *P*, en el vértice. El gradiente *Q* puede ser lineal o no lineal. La sección cónica creada por los planos al cortar el cono no es el patrón del ala, que se forma al proyectar la sección cónica sobre el ala. Los patrones mostrados son de complejidad creciente. De izquierda a derecha y de arriba abajo, las especies son *Morpho peleides*, *Smyrna blomfildia*, *Dichorragia nesimachus* y *Speyeria aphrodite*.

media de la celda alar. Patrones así pueden explicarse por el modelo de doble gradiente, si suponemos como mínimo que un gradiente no es lineal y que el gradiente de Q tiene una cresta o punto máximo que discurre a lo largo de la línea media, decreciendo simétricamente hacia las venas alares. Lo que este modelo sugiere es que, en la determinación de un patrón de coloración, intervienen, como mínimo, dos factores, distribuidos como los valores relativos de P y Q.

Una prueba indirecta de este modelo lo constituyen los experimentos publicados por Kühn en 1926. Sometió a pupas del ninfálido europeo *Argynnis paphia* a temperaturas inferiores a la de congelación durante el período de determinación del patrón de coloración. Aunque se desconocen los efectos fisiológicos de estos *shocks* térmicos, sí sabemos que conducen a un bloqueo manifiesto del patrón de coloración en diversos estadios del desarrollo. Podemos explicar mediante el modelo de dos gradientes la secuencia de patrones obtenida, si suponemos que, al establecerse por vez primera los gradientes en las etapas iniciales del desarrollo del patrón, existe un incremento progresivo del nivel de Q junto a una extensión gradual del gradiente cónico de P.

Para dar cuenta de los miles de patrones alares basados en el plano general ninfálido no se requieren más factores que los mentados. En consecuencia, los patrones propios de cada especie pueden surgir a través de alteraciones de sólo algunas variables durante el desarrollo: número y localización precisa de los focos; formas relativas de los dos gradientes, uno producido por el foco y el otro en algún lugar de la región proximal de la celda alar; elección de los pigmentos en que se expresa el patrón y magnitud de las señales posicionales propias de los dos gradientes que codifican para cada uno de los pigmentos. De todo este conjunto de factores sólo conocemos la existencia de los focos. El reducido grupo de investigadores que trabajamos en este campo estamos aún lejos de entender la naturaleza de los gradientes posicionales y los mecanismos a través de los cuales intervienen en la síntesis de pigmentos específicos de las escamas de los Lepidópteros. Sin embargo, disponemos ya de un marco conceptual en el que plantearnos cuestiones de interés, no sólo sobre la génesis de determinados patrones de coloración, sino también sobre la fisiología comparada y la evolución de los patrones biológicos en general.

Establecimiento ballenero vasco del siglo XVI en el Labrador

Los archivos y los restos arqueológicos encontrados en tierra y en el fondo del mar revelan una floreciente industria ballenera española en el Canadá, establecida el primer siglo después del descubrimiento del Nuevo Mundo

James A. Tuck y Robert Grenier

Los pescadores europeos empezaron a faenar en aguas de la costa oriental del Canadá a partir de 1497, año en que, al regreso del primer viaje conocido al golfo de San Lorenzo, John Cabot trajo la noticia de la existencia en aquellas aguas de bacalao en tal cantidad que podía recogerse a cestadas. Así, pues, no es extraño que, en poco tiempo, el lugar se convirtiera en la meca de los pescadores europeos y particularmente del golfo de Vizcaya.

En el siglo XVI, los pescadores del golfo de Vizcaya, los vascos principalmente, realizaban en aguas canadienses campañas de varios meses de duración, dedicándose no sólo a la pesca del bacalao, sino también a la caza de ballenas. Por esta razón, y debido a la larga permanencia en aquellos mares, se vieron obligados a fundar establecimientos o factorías en tierra, donde poder fundir el lardo o grasa y obtener el saín, o aceite de ballena, que envasaban en toneles y almacenaban hasta el momento de emprender el viaje de regreso a España, a principios de invierno. Los vascos cazaban las ballenas en un lugar que denominaban bahía Grande, y explotaban una industria muy floreciente de la que nada se sabía prácticamente hasta hace muy poco.

A principios de la década de 1970, y fruto de una investigación en archivos españoles, salieron a la luz las pesquerías de ballenas en bahía Grande, al mismo tiempo que los trabajos arqueológicos, estimulados por los documentos del siglo XVI, revelaron la existencia de los restos de un establecimiento pesquero vasco en la costa meridional del Labrador y condujeron al hallazgo de los restos de una nave, igualmente vasca, con un cargamento en bastante buen estado de conservación, en las proximidades de dicha factoría. Merece la pena detenerse en los trabajos de investigación efectuados en los archivos, todavía en curso, que propiciaron los descubrimientos, y examinar cuanto se ha encontrado en el establecimiento del Labrador en las cuatro últimas campañas arqueológicas. En tales campañas, el personal de la Universidad Memorial de Terranova se encarga de las excavaciones en tierra, en tanto que las submarinas las realiza la Unidad de Excavación Marina de Parques del Canadá. Sin embargo, desde el año pasado todos los trabajos se han convertido en un proyecto multidisciplinar, cuya culminación requerirá otras tres campañas, y en las que participan las instituciones siguientes: Parques del Canadá, Consejo de Investigación Canadiense de Ciencias Sociales y Humanas, Gobierno de Terranova y Labrador, Archivos Públicos del Canadá y el Instituto Canadiense de Conservación de los Museos Nacionales del Canadá.

La investigación en los archivos españoles la lleva Selma Huxley Barkham, experta en historia vasca, y su trabajo consiste en la traducción y análisis de millares de documentos notariales, prácticamente olvidados, tales como contratos, fletamentos, pleitos, pagarés y testamentos relacionados con las pesquerías de bahía Grande, y que se encuentran en diversos archivos vascos, principalmente en el de Oñate. Gracias a estos documentos han salido a la luz bastantes nombres vascos de lugares situados en la región de bahía Grande, revelando, al mismo tiempo, la organización de varias expediciones balleneras, algunas de las cuales tuvieron éxito, en tanto que otras terminaron en un rotundo fracaso. A una de estas últimas pertenecen, sin duda, los restos del buque naufragado y que encontraron los miembros de Parques del Canadá a partir de una relación de 1565 relativa a la pérdida de un buque, el *San Juan*, poco antes de emprender el viaje de regreso a España con un cargamento completo de saín o aceite de ballena.

Los estudios realizados hasta el momento por la señora Barkham evidencian los beneficios derivados de las expediciones balleneras en aguas canadienses durante el siglo XVI. Uno de los documentos relata que, en 1571, Doña Marina de Urançu aseguró la quinta parte del cargamento de aceite de ballena procedente del Nuevo Mundo en 2000 ducados, de lo que se desprende que el valor total del mismo debía ser de 10.000. Este importe a su vez equivalía al precio de dos galeones grandes comprados en Sevilla aquel mismo año. De ahí resulta que un armador o sociedad de armadores podía, en una sola campaña ballenera afortunada, recuperar el capital invertido en la adquisición de la nave y además obtener un pingüe beneficio.

Los mismos documentos nos proporcionan también una breve síntesis de la actividad de los balleneros. Sabemos que las naves zarpaban en dirección a bahía Grande en primavera. Un contrato de 1566 cita el caso de una salida muy temprana y que tuvo lugar el 25 de abril; pero normalmente los viajes se iniciaban algo más tarde. Una vez llegaban a su destino, los balleneros cazaban las ballenas arponeándolas desde chalupas o botes de unos cinco metros de eslora. La temporada de caza duraba hasta diciembre o principios de año, época en que los hielos les obligaban a regresar a España. A veces debían partir a toda prisa para escapar de los hielos, en particular cuando el frío se presentaba repentinamente, pues, de no hacerlo, tenían que invernar en el establecimiento. De hecho, los años de mediados de la década de 1570, y concretamente los inviernos de 1574-75 y

ENSENADA INTERIOR de la bahía Roja, en el Labrador, vista desde el norte, en primer término. Al fondo, detrás de la isla Saddle y de la otra más pequeña, que es la isla de Penney, se halla el estrecho de Belle Isle. Los balleneros vascos estuvieron aquí, a 3000 millas de España, en el siglo XVI.

PARTE CENTRAL DEL PECIO. Se trata probablemente del buque *San Juan*, hundido en 1565 y descubierto en 1978, frente a la isla Saddle. Se aprecian la sobrequilla, tres cuadernas, el forro del costado de estribor, toneles aplastados y lastre del buque, así como el dispositivo de referencia.



RESTOS DEL TEJADO descubierto en la costa norte de la isla Saddle, cuya investigación corre a cargo de arqueólogos submarinos de la Universidad Memorial de Terranova. Entre los restos aparecieron trozos de cerámica, cuchillos y otros objetos hechos en España en la segunda mitad del siglo XVI.

1576-77, fueron muy fríos; por tal motivo, según las investigaciones de Barkham, los balleneros quedaron bloqueados por los hielos y tuvieron que pasar el invierno en bahía Grande.

Durante la invernada de 1576-77 murieron varios hombres, la mayoría como consecuencia del frío o de la escasez de provisiones. Sin embargo, uno de ellos, por lo menos, fallecido en junio de 1577, pudo ser víctima del escorbuto o de otra enfermedad carencial. Se trataba de un tal Juan Martínez de Larrume, según se desprende de su testamento (el documento más antiguo de este tipo redactado en el Canadá). Pese a los riesgos propios de la caza de ballenas desde botes, la mayoría de los fallecimientos registrados en bahía Grande fueron debidos a los efectos de la invernada y no a accidentes de caza.

Las investigaciones de Barkham han proporcionado también mucha información sobre las técnicas empleadas en la caza de ballenas y el sistema de vida de los balleneros. La presa más codiciada era la especie *Balaena mysticetus*, es decir, la ballena de cabeza arqueada, que se alimenta de *krill*, y que los ingleses denominan *right whale* (ballena verdadera), probablemente porque era la más fácil de cazar: nadaba muy despacio y una vez muerta no se hundía, lo que hubiera representado su pérdida. La caza de ballenas se realizaba de forma parecida a como se describe en *Moby-Dick*: una vez arponeada y agotada se la remataba con la sangradera. (En el año 1566 el capitán Miguel de Cerain embarcó 100 arpones y 24 sangraderas con destino a bahía Grande; los primeros debían perderse más fácilmente que las segundas.) A continuación la ballena se remolcaba hasta la orilla, donde se encontraba el establecimiento y, una vez allí, era troceada y convertida en aceite. A veces, cuando el viento y la corriente eran en contra, la operación de remolque del mamífero resultaba sumamente difícil y en tales casos, según revela un documento, solían dejarlo sólidamente amarrado a cualquier punto de la costa hasta que las condiciones del tiempo experimentaban una mejoría. En 1575, Nicholás de la Torre fue acusado ante los tribunales por haber dejado que sus hombres se apoderaran y descuartizaran una ballena capturada por la gente de Joan López de Rey y que encontraron amarrada a una roca de la costa. El litigio duró veinte años y lo terminaron las viudas de los dos balleneros.

En lo concerniente a la vida diaria en los establecimientos de tierra, son de particular interés los datos relativos a las provisiones que se embarcaban con destino a bahía Grande. De momento, la señora Barkham ha encontrado dos listas de provisiones, y en las que figuran comestibles, principalmente tocino, bacalao salado (pese a que ese pescado se capturaba en aquellas aguas), sardinas, trigo, habichuelas, guisantes, aceite de oliva, semillas de mostaza, ajos, vino de jerez, sidra y gran cantidad de galleta. El equipo empleado en la caza de ballenas comprendía, además de chalupas, arpones y sangraderas, cabos de cáñamo de diferentes gruesos y cuchillos de formas diversas para descuartizar y trocear la grasa. También se mencionan las medicinas que administraba el barbero de a bordo, que hacía las veces de médico o cirujano, aunque sin detallarlas.

Una buena parte de los efectos importados eran materiales de construcción: clavos, tierra (probablemente arcilla) y tejas. La madera no se menciona, de modo que la empleada en las obras de la factoría debían obtenerla de los árboles del país, cortados con hachas igualmente traídas de España. (El capitán de Cerain llevó también 16 hachas a bahía Grande.) El modo como construían los edificios no lo refleja ningún documento, aunque había mu-



HUESOS DE BALLENA BLANQUECINOS cubren parte de la playa existente al oeste de la bahía Roja. Los arqueólogos contaron que procedían de 30 ballenas, como mínimo, de la especie *Balaena mysticetus*. En el fondo del mar, y frente a la isla de Penney, descubrieron también otro depósito similar.

cha demanda de tejas para sustituir las viejas y rotas; en 1563 Francisco de Florriaga las cargó en "cantidad suficiente para reparar el tejado de las cabañas" que habían sufrido las consecuencias de las inclemencias del tiempo durante el invierno anterior. Sólo existe una cita indirecta de la función de dichas edificaciones, y la proporciona el arponero Simón de Azcoitia, testigo en un juicio, al indicar que el sitio donde estuvo conversando con un compañero estaba situado "junto a la cabaña que habían construido y en donde fundían la grasa de las ballenas que mataban".

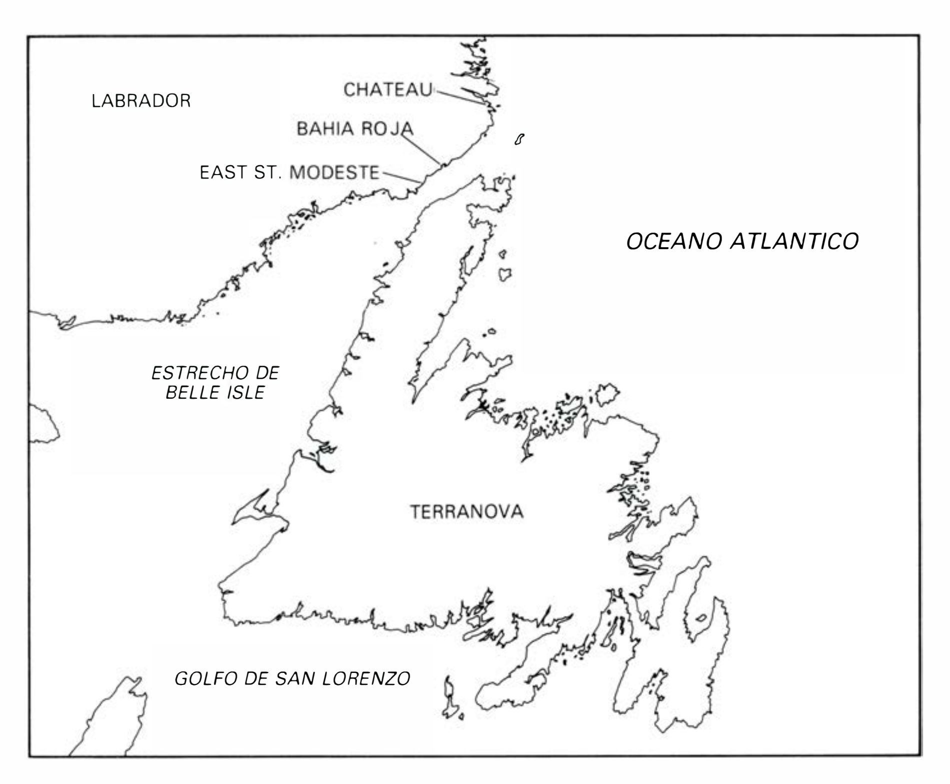
¿Dónde estaba bahía Grande? Tras estudiar los nombres geográficos que aparecen en los documentos, Barkham opina que los balleneros vascos bautizaron así al estrecho de Belle Isle, existente entre la parte septentrional de Terranova y la meridional de las costas de Quebec y Labrador, y que comunica el Atlántico con el golfo de San Lorenzo. Otro de los lugares que aparece en los documentos vascos es Samadet o Semadet, nombre que aplicaron al parecer a East St. Modeste, un antiguo establecimiento situado en la costa del Labrador, a pocas millas al oeste de la bahía Roja que está en el citado estrecho. Otro lugar mencionado por los vascos es Xateau o Chateo, que coincide posiblemente con el actual de Chateau, en la costa del Labrador. Hay también muchas citas de un tercer lugar, denominado indistintamente Butus, Buytes, Boytus y Buitres, y que al parecer se trataba de uno de los principales establecimientos vascos.

La bahía Roja es uno de los mejores puertos naturales del estrecho y se encuentra en la costa del Labrador. Consiste en una ensenada pequeña, que se abre en el interior de otra mayor, rodeada por un acantilado de 90 a 150 metros de altura, y cuya bocana está protegida por una isla baja con una colina en cada extremo, que los hidrógrafos del Almirantazgo Británico bautizaron con toda propiedad con el nombre de isla Saddle y que podríamos traducir por isla de la Silla de Montar.

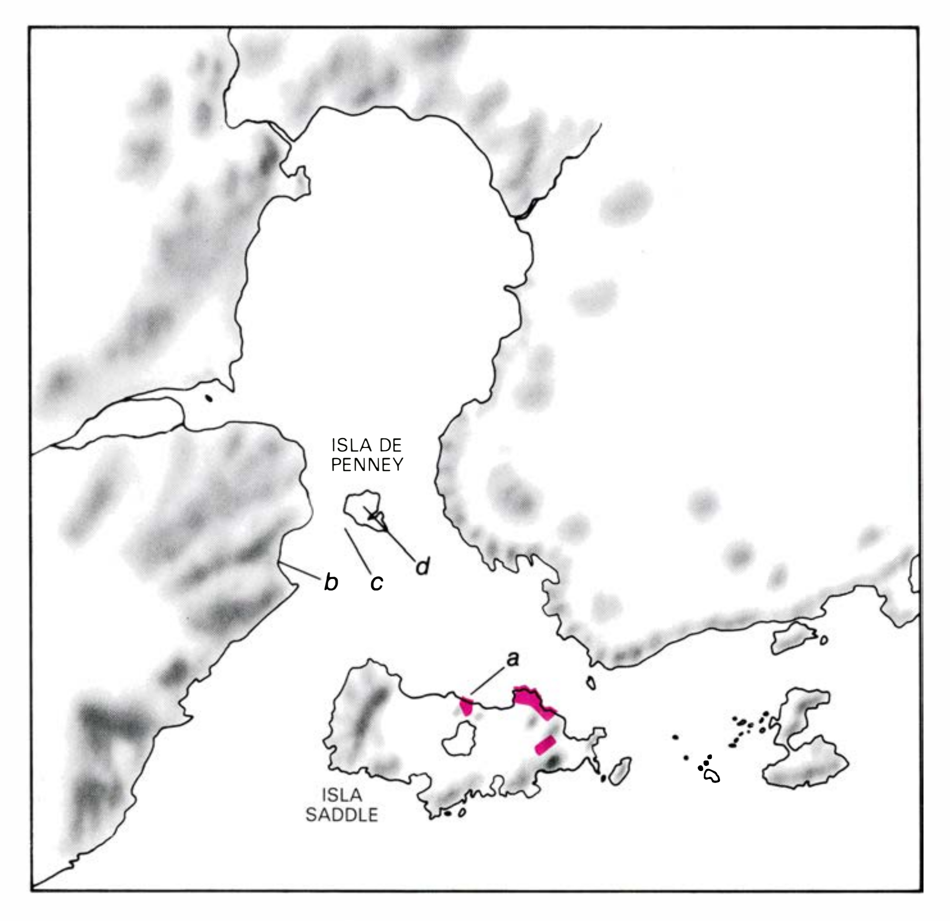
Por todo ello parecía muy acertado hacer una visita de inspección a la bahía Roja. Así, pues, en 1977, la señora Barkham y uno de nosotros, Tuck, acompañados por Walter Kenyon, representante del Museo Real de Ontario, y otros más, visitamos la pequeña localidad allí existente. Una vez allí, descubrimos que el suelo y las playas estaban llenos de pedazos de tejas rojas. Lo mismo ocurría en las playas de la isla Saddle, donde encontramos además varias paredes de piedra con manchas de una sustancia negra solidificada y cuyo análisis posterior reveló tratarse de grasa animal quemada. Por tanto, llegamos a la conclusión de que los pedazos de tejas rojas procedían de las empleadas por los vascos para techar las cabañas, y que la sustancia negra en las paredes de piedra indicaba el lugar donde se fundía la grasa de ballena para convertirla en aceite.

Nuestros descubrimientos eran suficientemente prometedores para buscar los fondos necesarios que nos permitieran efectuar unas excavaciones de un mes de duración, el año siguiente. Los trabajos de la primera campaña los concentramos en tres zonas de la isla Saddle que aparentemente no habían experimentado ningún cambio desde el momento en que fueron abandonadas, unos 400 años antes. Era el primero de ellos una pequeña terraza, desde la que se divisaba la bahía; descubrimos allí las ruinas de un tejado. Junto a ellas, encontramos restos de cerámica y cristal, así como objetos de hierro y utensilios, lo que permitió fechar la obra en la segunda mitad del siglo XVI.

El segundo lugar de excavación estaba situado en una marisma, cuyo grado de humedad era muy adecuado y favorecía la conservación de la materia or-



LOS BALLENEROS VASCOS DEL SIGLO XVI conocían con el nombre de bahía Grande el estrecho de Belle Isle, formado por un entrante del golfo de San Lorenzo y que separa Terranova de las costas de Quebec y Labrador. El actual East St. Modeste pudo muy bien ser el establecimiento denominado por los vascos Semadet; Chateau, el que conocían por Xateau, y el asentamiento de bahía Roja, Butus.



DETALLE DE LA BAHIA ROJA. La mayor parte de las excavaciones realizadas han tenido lugar en las zonas coloreadas de la isla Saddle; el buque hundido se encuentra en *a*. El depósito de huesos ilustrado en la página precedente se encuentra en *b*, al noroeste de la isla Saddle. Frente a la isla de Penney existen otros vertederos de huesos, así como los restos de una embarcación pequeña hundida (*c*). Los hornos para fundir el lardo o grasa están en *d*. Dicha ensenada está rodeada por un acantilado, de 90-150 metros.

gánica. Al abrir la primera zanja descubrimos una gran cantidad de virutas de madera, lo que nos indicaba que trabajaron troncos de árboles de la zona. Por el suelo y sin enterrar encontramos también trozos de aros o cinchos de tonel, varios pedazos de piel y unos fragmentos de madera de roble, árbol que no se da en el Labrador.

El tercero era el lugar donde habíamos encontrado los residuos de grasa animal, y que pensamos se trataba del horno. Al iniciar las excavaciones descubrimos gran cantidad de tejas, enterradas bajo una capa de arena que el oleaje había acumulado allí. El muro de piedra ennegrecido que atrajo nuestra atención durante la primera visita estaba situado en el medio de lo que en su día había sido un edificio muy espacioso.

Al término de la campaña de excavaciones en la bahía Roja el grupo de trabajo de la Universidad Memorial quedó plenamente satisfecho de los resultados, aunque el descubrimiento más espectacular de 1978 lo hicieron poco después los submarinistas de la institución Parques del Canadá: al hallar, sólo a 30 metros de la orilla septentrional de la isla Saddle, los restos de un buque hundido, con un cargamento de toneles de madera, cuyos aros y cinchos eran exactamente iguales a los encontrados antes en tierra. Además, el buque encajaba perfectamente con la mención de un naufragio registrado en los archivos vascos. De hecho, los datos aportados por Barkham sobre el buque naufragado, el *San Juan*, fueron los que indujeron al personal de Parques del Canadá a realizar la investigación submarina.

En la campaña siguiente, además de excavar en diversos lugares de la isla Saddle, los arqueólogos, una vez levantados los planos y fotografiados los pedazos de material del tejado derrumbado que habíamos descubierto en 1978 y que fue designado como Zona *A*, los sacaron de sitio. La superficie cubierta por los restos era de 14 metros de largo por 8 de ancho, de modo que la habitación, parte de la cual había sido construida en piedra, debía tener unas medidas ligeramente menores. En la isla descubrimos también otras paredes hechas a base de una parte central de grava y guijarros, recubierta por ambos lados con piedras de mayor tamaño, cuidadosamente encajadas. Aquí, no obstante, quedaba sólo el centro de grava, aunque permitía deducir el lugar donde estaba situada la cara frontal de la pared norte. La pared trasera no existía, pues habían utilizado en su lugar un

muro vertical de roca. (El aprovechamiento de elementos naturales en las obras de construcción es muy usual entre los vascos.)

De las paredes de levante y poniente, mucho más cortas, en el supuesto de que realmente hubieran existido, no queda ninguna traza. Es posible que se tratara de un cobertizo abierto por ambos extremos o que las paredes fueran de madera, en cuyo caso la acidez del suelo de la isla provocaría su rápida descomposición. La estructura del techo debía ser de madera, según se desprende del hecho de haber encontrado varios centenares de clavos forjados a mano. De todos modos es imposible determinar si el tejado llegaba hasta el canto del muro o por el contrario presentaba un cierto saliente, pues ambos estilos son muy comunes en las construcciones vascas.

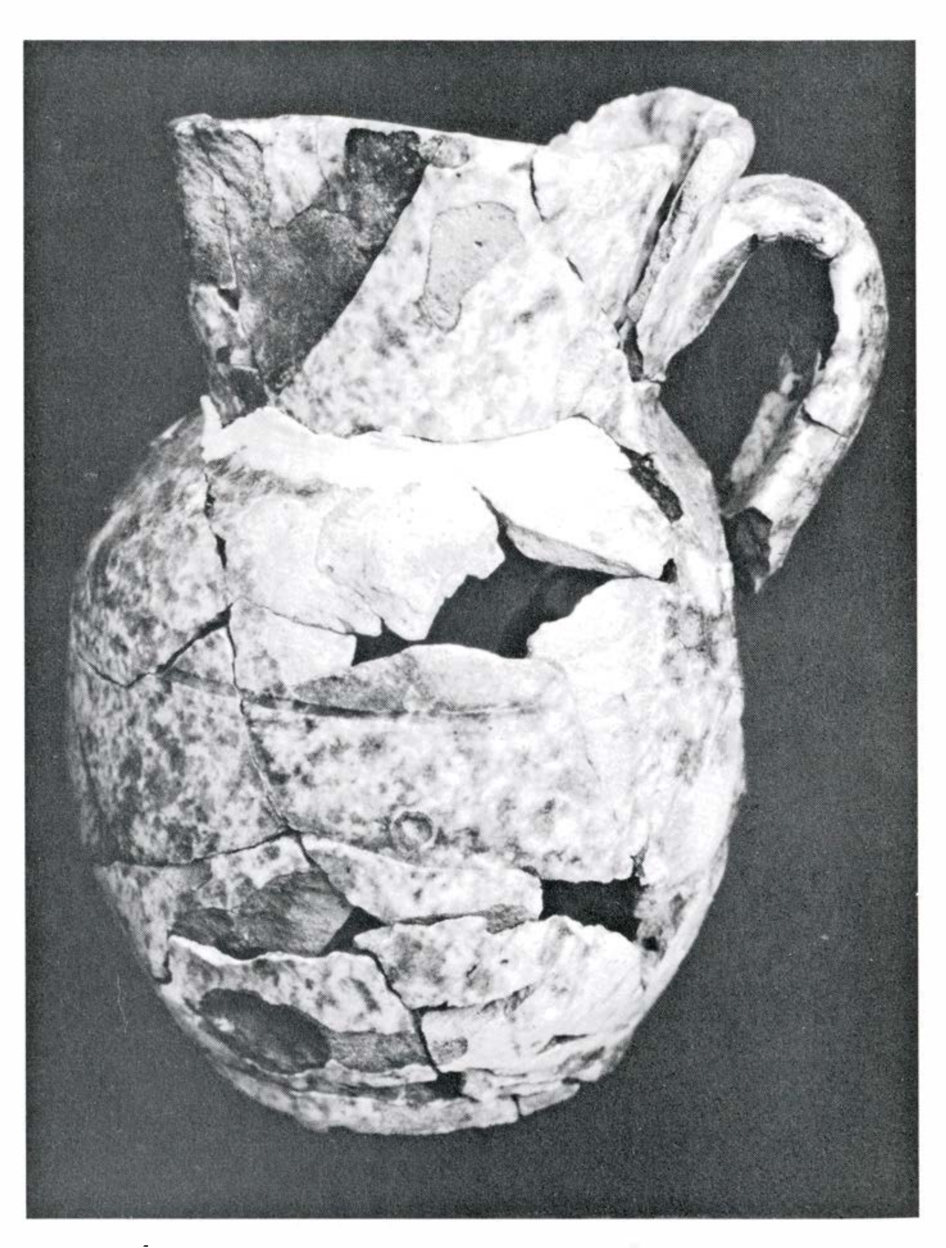
Debajo del tejado se descubrieron dos pequeños hogares, que fueron excavados en 1980, y se encontraron restos de carbón vegetal, huesos de animales quemados, pedazos de cerámica, perdigones de plomo y desperdicios procedentes de la fundición del plomo. De ahí se desprende que, aparte de caldear el ambiente y fundir la grasa, los hogares se utilizaban para cocinar y fundir proyectiles.

Dentro y fuera del edificio, principalmente en el extremo oeste, donde al parecer vertían los desperdicios, encontramos muchos utensilios a flor de tierra. Los más abundantes eran trozos de cerámica, cuyo número total ascendió a más de 2000. Gran parte de ellos procedían de vasijas, cacharros de cocina, jarras, escudillas y platos. Las tapaderas de las jarras y escudillas más pequeñas estaban decoradas en azul, verde u oro, sobre fondo blanco acristalado. Todos estos utensilios de cerámica (que los ingleses designan como *majolica* debido a que la isla de Mallorca es uno de los grandes centros de producción) eran de uso muy corriente en España y constituían uno de los principales productos de exportación del país en el siglo XVI. Los utensilios de cristal eran menos abundantes: sólo se encontraron pedazos de una copa de vino y de una jarra grande decorada con un relieve de espoletas.

Además de los clavos y puntas de la estructura de madera, los arqueólogos recuperaron también muchos utensilios de metal. Entre ellos figuraban varios cuchillos de hierro, uno de los cuales llevaba un mango de madera tallada con un pomo de bronce en forma de corona. Del análisis de los hogares, cerá-



DUELAS Y TABLAS de los fondos de los toneles encontrados en la bodega del buque hundido frente a la isla Saddle se conservan en perfecto estado; nadie diría que tienen 400 años de edad. Las señales dejadas en las duelas por los aros o cinchos se distinguen todavía perfectamente. Las duelas eran de roble y haya.



UNA PEQUEÑA JARRA que forma parte de los numerosos objetos de cerámica española encontrados durante las excavaciones del establecimiento costeño donde trabajaban los toneleros españoles.

mica y cuchillos se desprende que el recinto había servido de vivienda. Esta impresión se ve reforzada por el hallazgo de otros restos, como un rosario de madera, una moneda y una llave. Al término de la campaña de 1979, todo parecía indicar que en él habían vivido personas de cierta categoría.

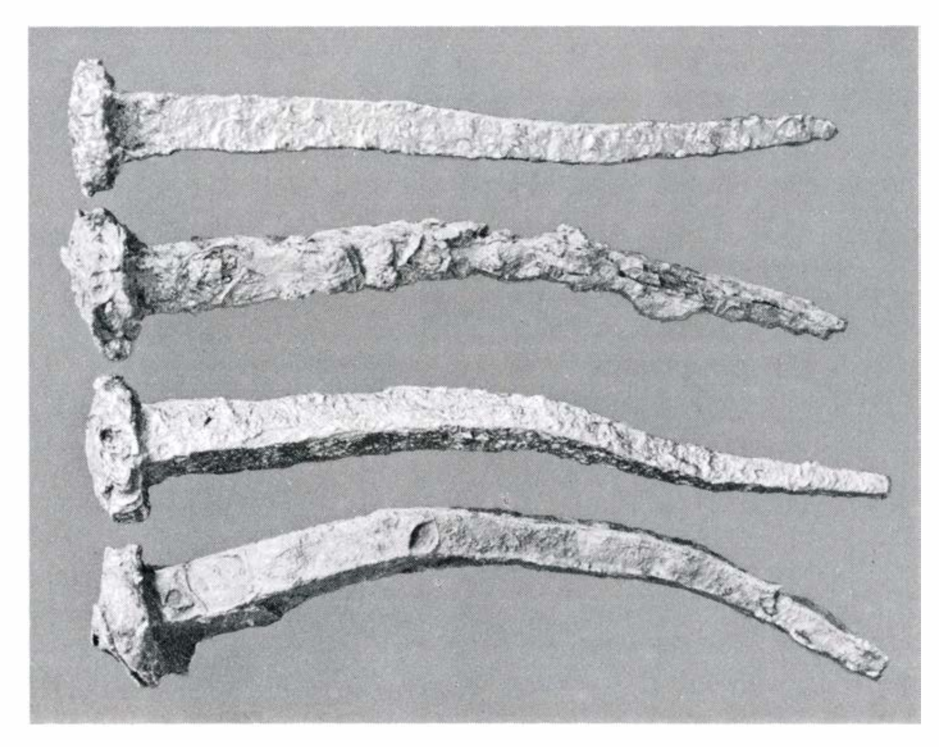
La temporada siguiente, las excavaciones realizadas en un lugar próximo de la misma isla, en la llamada Zona *E*, modificaron esta opinión. En este punto, de características muy similares a las de la Zona *A*, se descubrió una estructura y gran número de utensilios, entre los que figuraban cacharros de cocina, pedazos de cristal y otros objetos de uso doméstico, de idéntica factura a los anteriores. Entre ellos, sin embargo, figuraban varias herramientas, lo que indicaba que sus ocupantes eran toneleros. Dichas herramientas eran dos hachas, una de las cuales estaba tan bien conservada que cortaba todavía, una cuña o chasa para colocar los aros o cinchos y varios tornillos de gatos o prensas de armar toneles. Todos estos hallazgos permitieron aclarar la función de dos pedazos de metal encontrados anteriormente en la Zona *A*, y que no pudieron identificarse hasta que el Instituto Canadiense de Conservación los fotografió con rayos X. Uno de ellos era un pedazo de cuchilla de un cepillo de carpintero y el otro un diente de sierra. Por todo ello es muy probable que los ocupantes de ambas habitaciones estuvieran dedicados a trabajos de carpintería.

Las posibles dudas se disiparon inmediatamente, al menos en lo que concierne a la Zona *E*. Al lado de la pared norte el agua se había infiltrado por los huecos existentes entre las piedras del suelo, arrastrando algunos desperdicios, de modo que al excavar allí encontramos restos de las virutas producidas durante la construcción de toneles, y cuyo estado de conservación parecía indicar que tuvieran sólo unos pocos años de antigüedad, en vez de siglos. Cerca de allí había también piezas de toneles que debieron ser rechazadas a causa de algún defecto en la madera o en su construcción. Las duelas de roble y haya y las piezas de los fondos de los toneles coincidían con las procedentes del buque naufragado, al igual que ocurría con los aros o cinchos de aliso y mimbre con que se aseguraban. Descubrimos también pedazos de madera, sobrantes del redondeo de los extremos de las tablas que forman los fondos de los toneles, aros cortados y trozos de sauce con muchos nudos o excesivamente pequeños para su empleo, así como virutas y serrín. Por tanto, es indudable que la Zona *E* servía de residencia y taller de construcción de toneles a la vez; la similitud entre ésta y la habitación de la Zona *A* permitía deducir que ambas estaban destinadas a idéntica función. De todos modos, nuestras conclusiones de 1979 quizá tuvieran algún punto de verdad: tal vez los toneleros, cuyos toneles eran imprescindibles para el transporte del aceite de ballena, disfrutaran de una posición privilegiada en las factorías de tierra.

Los hornos donde se fundía la grasa o lardo estaban situados entre los talleres de construcción de toneles y la playa. Uno de aquellos, cuyos restos de la obra de piedra se elevan todavía hasta la altura de un metro, fueron excavados en 1980. Dicha obra, de unos 10 metros de largo y que discurre en dirección paralela a la orilla, constituía la base del fogón. Los calderos iban en el lado del mar de la misma, en tanto que al otro lado había una zona de trabajo. La obra en sí estaba construida de forma muy cuidadosa, a base de piedra del lugar y lastre del buque, unida con mortero de arcilla, probablemente la "tierra" que aparece en las relaciones de materiales importados. Una vez la arcilla quedó impregnada de aceite de ballena formó una especie de argamasa que cementó las piedras entre sí.

¿Cómo derretían los vascos el lardo o grasa? No hay ninguna descripción que explique el sistema que empleaban, pero sí sabemos cómo lo hacían los ingleses y holandeses que se dedicaban a la caza de ballenas en la isla de Spitsbergen, en el Océano Artico, que lo aprendieron de ellos, y cuyos relatos han llegado hasta nosotros. En Spitsbergen las ballenas muertas se remolcaban con los botes o desde tierra, mediante el empleo de aparejos, hasta el lugar donde eran descuartizadas. (En la playa de poco fondo, situada frente a los hornos descubiertos en la isla Saddle, existen los restos de una especie de varaderos, construidos a base de madera ensamblada y piedra de lastre, donde probablemente troceaban las ballenas.) Los pedazos de lardo o grasa, debidamente troceados, se metían en toneles y a continuación éstos se subían por una rampa hasta los calderos situados en el lado del mar de la pared del horno. En Spitsbergen, los calderos eran de cobre con una capacidad de 48 a 54 galones cada uno (unos 182 a 204 litros, aproximadamente), equivalentes a un barril de grasa troceada, e iban dispuestos en línea sobre una obra de ladrillo que servía no sólo de soporte sino también de hogar. La parte superior de los mencionados calderos se encontraba a un par de metros sobre el suelo. (En el lado del mar de la pared de la isla Saddle se aprecian los restos de seis hogares de piedra y arcilla.)

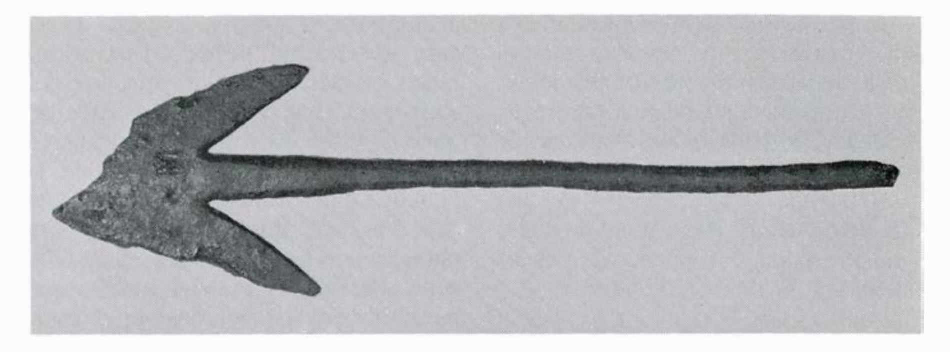
El capitán Thomas Best, buen conocedor de los establecimientos de Spitsbergen, informó que se solía colocar una chalupa o bote viejo lleno de agua al otro lado de la pared donde estaban los hornos. Dos hombres, subidos a



CLAVOS DE HIERRO FORJADO, de 17 a 18 centímetros de largo, hallados entre centenares de puntas del mismo material usadas por los vascos en la construcción del establecimiento de la isla Saddle.

una plataforma o andamio dispuesto al efecto junto a ella se encargaban de sacar el aceite o saín de los calderos y verterlo dentro del bote, donde se enfriaba y purificaba, pues las impurezas se precipitaban al fondo mientras el aceite o saín quedaba flotando en el agua.

En el lado de tierra de la pared de la isla Saddle se encontraron restos de una plataforma o andamio similar, conservado por la saturación de aceite. Dicho andamio estaba hecho principalmente de madera local, aunque algunas tablas de roble con agujeros evidenciaban que los constructores vascos habían empleado también maderas de botes o barcos viejos. Por otra parte, en vez de enfriar el aceite en un bote viejo lleno de agua, los encargados de vaciar los calderos lo vertían en unas cubas o toneles cortados por la mitad, de unos 91 litros de capacidad, que ponían sobre el mismo andamio. Las duelas, fondos y tablas de roble o abeto, así como los aros o cinchos encontrados entre la madera que formaba dicho andamio, eran exactamente iguales a las piezas de los toneles existentes a bordo del buque hundido. Junto a la pared encontramos también tejas y restos de una estructura de madera que evidenciaban que andamio y calderos estaban bajo cubierto y protegidos de los elementos.



PUNTA DE ARPON, que originariamente tenía 40 centímetros de largo, encontrada en la bahía Roja, perteneciente a los balleneros españoles. En ella se aprecia una letra *M*, que probablemente es la marca del propietario, y que parece un monograma vasco del siglo XVI. Los arpones, sangraderas y cuchillos de trocear las piezas cobradas figuran en las listas de las provisiones embarcadas con destino a Butus.

El documento que suscitó la búsqueda del buque hundido en Saddle corresponde a un proceso judicial. De él se deduce que, a fines de 1565, el buque *San Juan*, con un cargamento completo de aceite o saín de ballena, estaba haciendo los preparativos para iniciar el viaje de retorno a España cuando se perdió al "faltar el cable". Esta frase sugiere que una inesperada tormenta debió romper la gúmena del ancla del buque y arrastrarlo hasta la orilla, donde se hundió. En 1978, los submarinistas de la Unidad de Excavación Marina descubrieron, al norte de la isla Saddle, y a 10 metros de profundidad, los extremos de unos maderos que sobresalían del lodo del fondo. Una vez inspeccionados, los componentes del grupo llegaron a la conclusión de que pertenecían a los restos de un buque y a su regreso a tierra lo primero que vieron fue la zanja que los arqueólogos de la Universidad Memorial habían abierto en uno de los hornos del establecimiento.

El hecho de que el buque estuviera tan cerca de la factoría vasca permitía suponer que se trataba de un buque también vasco y probablemente del *San Juan*. Los trabajos preliminares de la Unidad de Excavación Marina vinieron a reforzar esta suposición. La mayor parte del maderamen del buque era de roble, lo cual imposibilitaba que se tratara de una nave de la misma época o posterior construida en el país. Además, en el buque se encontraron gran cantidad de duelas de roble y haya, así como tablas de los fondos de los toneles y aros o cinchos de hierro exactamente iguales a los hallados en la marisma de la isla. Por todo ello, se tratara o no del *San Juan*, lo cierto es que era un buque relacionado con los balleneros de la bahía Roja del siglo XVI.

A partir de 1978, los arqueólogos de los Parques del Canadá, con el apoyo del gobierno de Terranova y Labrador, han proseguido las excavaciones en la bahía Roja concentrando sus esfuerzos en dos objetivos principales: la excavación del naufragio en sí, operación que incluye el dragado de un canal desde la popa del mismo hasta la orilla, y la inspección general de la bahía. Esta inspección, realizada por los dos grupos de arqueólogos, condujo al descubrimiento de un lugar lleno de cráneos, vértebras, costillas y demás huesos de ballena. Igualmente encontraron los restos de dos hornos más, por lo menos, situados en la parte noroccidental de la isla de Penney.

Aparte de los huesos de ballena, la única evidencia visible, o producto de la excavación, de la actividad vasca en la isla de Penney eran unos cuantos fragmentos de tejas rojas. Frente a esta isla los submarinistas encontraron también otros dos depósitos de huesos de ballena, lo que parece indicar que una vez descuartizados los mamíferos y extraído el lardo o grasa, que se llevaba a tierra para convertirlo en saín o aceite, dejaban que los restos se hundieran por sí mismos.

Los submarinistas descubrieron, además, frente al extremo occidental de la isla de Penney, los restos de otro buque hundido, de menor tamaño al encontrado en la isla Saddle, y cuya construcción y clavazón parece indicar que es mucho más antiguo, aunque sólo los estudios posteriores nos revelarán si corresponde o no a la misma época del establecimiento vasco de la bahía Roja.

Además de la excavación iniciada en tierra se hicieron otras tres bajo el agua para determinar los límites del naufragio. Una primera hacia el norte, dirección que coincide con la proa del buque, y las dos restantes hacia el este y oeste, es decir, hacia babor y estribor del mismo. El buque descansa en el fondo sobre su costado de estribor; por tal motivo, los restos estructurales y demás efectos y objetos se encontraban bastante dispersos hacia ese lado. El fondo de la bahía es de lodo, piedra y conchuela. Al excavar en dirección a tierra el grupo de submarinistas descubrieron una capa de astillas de madera exactamente iguales a las encontradas a millares en tierra y que procedían del trabajo de la madera del lugar. Las astillas habían ido al fondo, en vez de flotar y ser arrastradas por las aguas; cabe suponer, pues, que se trataba de un vertedero de desperdicios y que previamente habían quedado empapados de agua en el lugar de construcción.

Sobre la capa de astillas los submarinistas encontraron otra de material orgánico, que en algunos lugares tenía un espesor de 20 centímetros, formado por huesos de aves y mamíferos terrestres, así como de ballenas, y gran cantidad de restos de bacalao, cabezas y vértebras sobre todo. Este hecho, unido a la ausencia de otras espinas, parece indicar que los vascos abrían el pescado y, una vez salado, lo ponían a secar. Esta hipótesis permite explicar algo que preocupaba al grupo de la Universidad Memorial desde su primera visita a la isla Saddle: el hallazgo de una serie de espacios circulares de unos dos metros de diámetro, cuyo suelo estaba recubierto de grava y piedra, y que utilizaban probablemente para apilar el pescado seco. En el depósito de huesos encontrados en tierra, los de ballena y

de bacalao aparecían mezclados, lo que indica que los vascos pescaban y cazaban ambas especies simultáneamente. En este mismo depósito se encontraron también los consabidos pedazos de tejas, así como restos de cerámica, trozos de jarcia, varios zapatos de piel muy bien conservados y parte de uno o dos botes pequeños empleados por los balleneros.

El principal objetivo de la investigación submarina es el propio buque naufragado. Esto no se debe solamente a la abundancia de restos y a su excelente estado de conservación, sino también a que de la madera del mismo es posible obtener una información única sobre la construcción de buques mercantes del siglo XVI. De acuerdo con los documentos estudiados por Barkham, el *San Juan*, al hundirse en aguas de escasa profundidad, permitió a sus tripulantes extraer de él la mayor parte del cargamento, así como algunas vituallas para el viaje de retorno a España y efectos personales. Por consiguiente, si el buque que naufragó en la isla Saddle es el *San Juan*, al hundirse, la parte alta del mismo debió quedar por encima del nivel del agua, y es muy posible que permaneciera así durante varias estaciones. En cualquier caso, los efectos del hielo invernal debieron acelerar su destrucción. El fango y los desperdicios, incluidos los pedazos del buque en sí, al amontonarse a ambos lados del casco, formaron un elemento protector que, unido a las bajas temperaturas del Labrador, ha conservado tanto los restos del naufragio como la parte de la carga que quedó a bordo.

Hasta el presente, los submarinistas han excavado una serie de agujeros cuadrados de dos metros de lado, la mayoría de ellos en la parte central del naufragio. Gracias a ello ha quedado al descubierto toda la parte del buque situada bajo la flotación, en el sentido de la manga, a proa y popa de la supuesta posición del palo mayor, así como las cabezas de proa y de popa del casco. Desde el momento en que iniciaron los trabajos, los submarinistas encontraron un buen número de restos de toneles de saín o aceite de ballena aplastados, aunque prácticamente completos. Algunos aparecían dispersos al azar, en tanto que otros mantenían la misma posición en que iban estibados, es decir, tendidos y apoyados sobre el plan del buque, con los extremos calzados con piedras de lastre y tacos de madera a fin



PARED DE PIEDRAS Y CASCOTES construida paralelamente a la orilla de Saddle. Los calderos empleados para fundir la grasa iban en el lado del muro que daba al mar. La zanja existente en el lado de tierra contiene fragmentos del andamio o plataforma, donde se situaban los hombres para sacar el saín.

de impedir que se movieran. Los toneles de aceite tenían un aforo de unos 204 litros. Al término de la campaña de 1980 los arqueólogos submarinos habían recuperado e inventariado más de 10.000 piezas de tonel.

Como la quilla del buque tiene unos 15 metros de eslora, es posible calcular que su capacidad de carga oscilaba alrededor de las 300 toneladas. Por otro lado, la forma de la proa y popa del buque, la suave curvatura de la roda y la disposición del espejo coinciden plenamente con los elementos correspondientes a la arquitectura naval del siglo XVI. En cambio, en algunos aspectos se aprecia una tecnología mucho más antigua. Así, por ejemplo, el extremo popel de la quilla se levanta formando un codillo al que va unido el codaste, sistema muy similar al empleado en la construcción de las naves mercantes medievales conocidas como cocas. (Se ha descubierto recientemente una en el puerto de Bremen.)

Salvo las piezas de toneles, los demás objetos encontrados entre los restos del naufragio son bastante escasos. A unos 22 metros al este del casco se descubrió un ancla, actualmente sometida a tratamiento para combatir los efectos de la inmersión. También se han recuperado el cabrestante, que es uno de los pocos conocidos del siglo XVI, y algunos motones del aparejo. Además, se han localizado dos pedazos de tubo de la bomba de achique, aunque no se han extraído aún. Posiblemente, el objeto más interesante de todos sea un falconete de hierro forjado. Este tipo de cañón de retrocarga, montado en la tapa de regala del buque, lo denominaban comúnmente los de habla inglesa "asesino", por cuanto su manejo solía producir tanto daño a los que disparaban como a los disparados. En cualquier caso hay otra incógnita: el cañón estaba cargado. Uno querría saber por qué.

A medida que avanzan los trabajos de excavación, y una vez fijada exactamente su situación en el hallazgo, las piezas de madera del buque se sacan a tierra, donde se dibujan y miden. De este modo, a partir de los datos obtenidos será posible construir un modelo reducido del buque en el estado en que se encuentra en el fondo. Luego, una vez terminada la excavación, con los restos se procederá a la reconstrucción del casco, lo que permitirá contemplar el aspecto que ofrecía. Todos estos trabajos aumentarán notablemente nuestros conocimientos sobre el proyecto y tecnología utilizados en la construcción de los buques de altura que llevaron al Nuevo Mundo a los exploradores, colonizadores y pescadores del siglo XVI.

El revestimiento del intestino delgado

Considerado, largo tiempo, un tejido pasivo, la investigación reciente ha puesto de manifiesto que las células del revestimiento intestinal están cubiertas por una membrana que digiere activamente los alimentos y los manda así a la sangre

Florence Moog

Si nos guiáramos por la apariencia inmediata diríamos que los alimentos entran en el organismo a través de la boca. La verdad es que no pasan a formar parte del mismo hasta después de haber cruzado el revestimiento intestinal para penetrar en la corriente sanguínea. La mayoría de los organismos superiores, mamíferos incluidos, muestran una disposición topológica consistente en dos cilindros concéntricos. El tubo digestivo se extiende desde la boca hasta el ano, y la masa corporal se dispone entre ese tubo estrecho y otro más amplio constituido por la piel. El interior del tracto gastrointestinal está, pues, más estrechamente relacionado con el medio ambiente que con los músculos, nervios, esqueleto y vasos sanguíneos del medio interno del organismo. El fenómeno fundamental de la asimilación de los alimentos consiste en atravesar la barrera existente entre la luz intestinal y el medio interno del organismo.

Se consideró en otro tiempo que la mayor parte de la digestión ocurría en la luz intestinal y que el recubrimiento del intestino se comportaba simplemente como una membrana pasiva. Una vez digeridos, los alimentos difundirían pasivamente sin la intervención activa de las células de la pared intestinal. Las investigaciones recientes de la fisiología presentan una imagen distinta de la función del revestimiento intestinal. Sabemos ya que la superficie interna del intestino desempeña un papel activo tanto en la digestión como en la absorción de las sustancias nutritivas. Las células del recubrimiento intestinal poseen enzimas en su membrana que transforman azúcares complejos en azúcares simples y descomponen los péptidos producidos en la digestión proteica en aminoácidos o en péptidos de pequeño tamaño, esto es, formados por unos cuantos aminoácidos.

Además, la membrana de las células del revestimiento intestinal dispone de una red de sistemas de transporte. Cuando la digestión ha terminado, los transportadores desplazan los azúcares, aminoácidos y péptidos de pequeño tamaño hacia el citoplasma de las células, desde donde pasarán finalmente a la sangre. Para cada molécula de azúcar o de aminoácido parecen existir varios sistemas de transporte; la mayoría de ellos no se conocen aún con claridad. Es del mayor interés para la fisiología del aparato digestivo descrifrar tales mecanismos de transporte. Con independencia de cuál de las explicaciones propuestas acabe por imponerse, la idea de una barrera intestinal que actúe como mera membrana pasiva es algo que pertenece ya al pasado. El recubrimiento intestinal es un tejido que digiere sustancias nutritivas, selecciona las que deben pasar a la sangre y acelera su transporte hacia el medio interno.

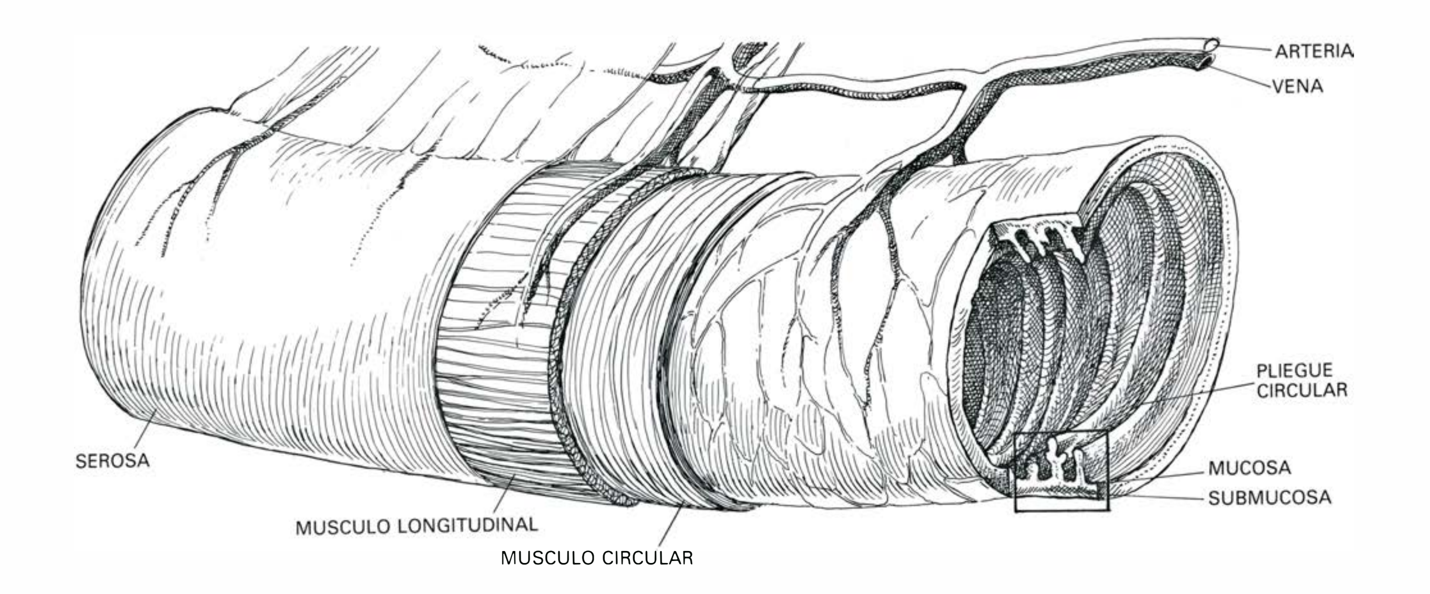
Los biólogos se han interesado desde hace tiempo por descifrar dónde y cómo tienen lugar la digestión y la absorción. Los experimentos llevados a cabo para aclarar tales mecanismos van desde los más primitivos, que ocasionalmente implicaban al propio tracto digestivo del investigador, hasta las investigaciones más sofisticadas sobre la estructura microscópica del revestimiento intestinal.

Hacia el año 1770, Lazzaro Spallanzani, de la Universidad de Pavía, llevó a cabo varios experimentos que ayudaron a interpretar la naturaleza de la digestión. Alimentaba su halcón doméstico con un trozo de carne atado a un cordel. Poco después de que el animal la hubiera deglutido, sacaba del estómago el trozo de carne tirando de la cuerda y comprobaba que había disminuido de tamaño y estaba parcialmente digerida. Spanllanzani ingirió también tres onzas de carne de pollo encerradas en un saco de ropa y observó, al expulsar el saco, que la mayor parte de la carne había desaparecido. En un experimento algo más elaborado, Spallanzani hizo deglutir a su halcón un tubo perforado que contenía una esponja. Extraía la esponja y obtenía, exprimiéndola, un jugo capaz de digerir carne.

Estos procedimientos le valieron a Spallanzani para determinar que la digestión empezaba en el estómago y que las sustancias que la llevaban a cabo podían atravesar la tela del saco con el que había realizado el experimento. Incluso llegó a aislar algunos de los líquidos digestivos merced a la esponja. A mediados del siglo XIX se sabía que la digestión continuaba en el intestino delgado, donde actuaban enzimas segregados por el páncreas. Si bien se conocía que el intestino delgado participaba en la digestión, se ignoraban aún los mecanismos a través de los cuales se absorbían las sustancias nutritivas.

PROYECCIONES DENOMINADAS VELLOSIDADES recubren la superficie interna del intestino delgado. En esta micrografía, obtenida en un microscopio electrónico de barrido del laboratorio del autor, en la Universidad de Washington, se muestra, aumentada 630 veces, una vellosidad intestinal del duodeno, o segmento superior del intestino delgado, de una rata. Las vellosidades, que dan un aspecto aterciopelado a la superficie del intestino, están cubiertas por una capa epitelial unicelular. No hace mucho se creía aún que la digestión ocurría únicamente en la luz intestinal, es decir, en el espacio interior del intestino, pero la fisiología acaba de demostrar que tanto la digestión como la absorción tienen lugar a través de estructuras moleculares ancladas en la membrana de las células epiteliales de la vellosidad.





ANATOMIA DEL INTESTINO DELGADO, que es similar en todos los mamíferos. La pared del intestino consta de cuatro capas concéntricas. Las tres externas, serosa, muscularis (formada por dos estratos musculares) y submucosa, rodean a la mucosa, que constituye la estructura más importante para la digestión y la absorción. Una serie de proyecciones dirigidas hacia la luz intestinal aumentan considerablemente la superficie de absorción. En los macromamíferos, hombre incluido, las proyecciones de mayor tamaño son los pliegues circulares. Tales pliegues, de ocho a 10 milímetros de altura, abarcan de la mitad a dos tercios de la circunferencia interna del intestino. Alcanzan su máximo desarrollo en el yeyuno, el segmento medio del intestino. Las vellosidades se proyectan hacia la luz intestinal partiendo de los pliegues y de la superficie que los rodea. En el intestino humano, las vellosidades varían en altura de 0,5 milímetros a 1,5 milímetros y se encuentran empaquetadas con una densidad que oscila entre 10 y 40 por milímetro cuadrado. En el interior de cada vellosidad existe una densa red de vasos sanguíneos y linfáticos hacia donde se transportan las sustancias nutritivas una vez finalizada la digestión. Entre las vellosidades hallamos unas depresiones, denominadas criptas, que descienden hasta el nivel de la muscularis. La ilustración es de C. Donner.

Hacia el año 1850 se vislumbró una posible explicación. Se descubrió que las moléculas disueltas podían difundir a través de una membrana desde una región de alta concentración hasta otra de menor concentración. Después de una comida, los alimentos pasarían por difusión pasiva de la luz intestinal, donde su concentración es elevada, a la sangre. Aunque la difusión pasiva podía explicar parte de las propiedades del revestimiento intestinal, en las dos décadas siguientes se puso de manifiesto que no servía para dar razón de todas ellas. Hacia 1870 se demostró que el agua podía atravesar la barrera intestinal de un segmento de intestino que contenía suero sanguíneo (la fracción líquida de la sangre). La difusión pasiva no podía explicar tal movimiento, puesto que la concentración de agua es prácticamente igual en la sangre y en el suero. Se había demostrado, además, que las moléculas de azúcar de seis átomos de carbono utilizadas en los procesos metabólicos se absorbían más rápidamente que aquellas que carecían de importancia metabólica. También se disponía de datos que indicaban una absorción selectiva de grasas y proteínas.

La mayor parte de la investigación ulterior sobre la fisiología del intestino delgado dedicó su esfuerzo a desvelar los mecanismos mediante los cuales las moléculas útiles para el organismo se digieren y transportan a través de la barrera intestinal con gran rapidez y en contra de un gradiente de concentración. Simultáneamente, se ha investigado la anatomía microscópica del recubrimiento intestinal. A partir de 1960, las dos líneas de trabajo convergen en el descubrimiento de estructuras ancladas en la superficie interna del intestino, que desempeñan importantes funciones digestivas y de absorción.

La estructura del intestino delgado, que responde a un patrón común en la mayoría de los mamíferos, refleja dos necesidades: el intestino debe presentar una superficie muy extensa para la absorción y ha de caber en el pequeño volumen de la cavidad abdominal. Para cumplir tales requisitos, forma asas y presenta varios tipos de proyecciones dirigidas hacia la luz intestinal que incrementan notablemente su superficie.

El intestino delgado se extiende desde la válvula pilórica del estómago hasta el inicio del intestino grueso. En una persona adulta viene a medir, por término medio, unos seis metros de longitud. A lo largo del intestino delgado pueden distinguirse tres secciones: el duodeno, que desciende unos 25 centímetros desde el estómago y está unido a la pared posterior de la cavidad abdominal por un tejido membranoso denominado mesenterio (el resto del intestino disfruta de una movilidad total); el yeyuno, segmento situado a continuación del duodeno y que ocupa la parte superior izquierda de la cavidad abdominal, y el íleon, en la parte inferior del abdomen. En los tres segmentos, la pared presenta la misma organización estratificada: tres capas exteriores, la serosa, la muscularis y la submucosa, rodean una capa interior denominada mucosa, que es la más importante desde el punto de vista de la digestión y absorción de los alimentos.

En los macromamíferos, el hombre incluido, existen tres tipos de proyecciones dirigidas desde la mucosa hacia la luz intestinal. En orden de tamaño decreciente y de mayor importancia fisiológica se suceden los pliegues circulares, las vellosidades (villi) y las microvellosidades (microvilli). Los pliegues circulares son crestas de ocho a 10 milímetros de altura que abarcan de la mitad a dos tercios de la circunferencia interna del intestino. Los pliegues comienzan cerca del extremo superior del duodeno y terminan hacia la mitad del íleon. Están más desarrollados en la última porción del duodeno y en el segmento inicial del yeyuno. Los pliegues aumentan la superficie intestinal hasta tres veces.

El intestino de los micromamíferos carece de los citados pliegues y posee únicamente vellosidades y microvellosidades. Las vellosidades, o villi, se proyectan a la luz intestinal desde los pliegues y desde la superficie mucosa que los rodea. Su densidad está comprendida entre diez y cuarenta por milímetro cuadrado y, a simple vista, dan a la mucosa una apariencia aterciopelada. Las vellosidades intestinales tienen una altura comprendida entre 0,5 y 1,5 milímetros. Entre las vellosidades existen depresiones, denominadas criptas, que penetran hasta la muscularis. En el interior de cada vellosidad se encuentra una red de vasos sanguíneos y linfáticos sostenida por una trama de tejido fibroso laxo. Precisamente hacia esos dos sistemas que ocupan la parte central de la vellosidad se transportan las sustancias nutritivas después de su digestión.

La vellosidad intestinal está cubierta por un epitelio unicelular. El epitelio, a su vez, está formado por dos tipos principales de células: las células caliciformes, que segregan un moco que recubre y protege la pared intestinal, y las células que realizan la absorción, implicadas en funciones de digestión y transporte. Entre las células epiteliales que recubren las criptas y las que recubren la vellosidad intestinal no hay solución de continuidad. No obstante, las primeras difieren de estas últimas en que son más cortas, aproximadamente cúbicas y en que algunos detalles de su superficie están menos desarrollados. Además, en las células de las criptas, pero nunca en las de la vellosidad, pueden observarse figuras mitóticas características de la división celular.

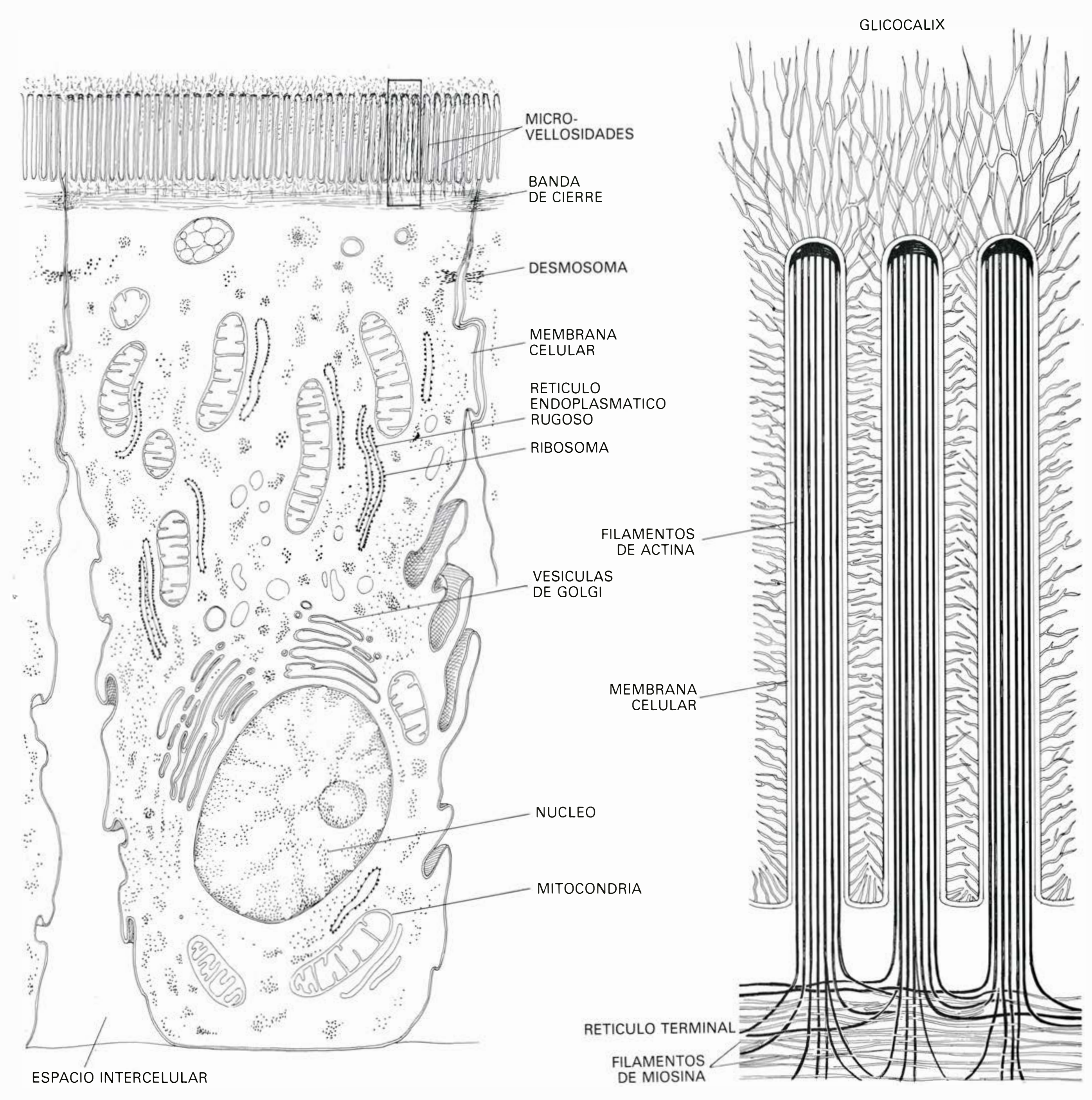
Las observaciones llevadas a cabo con el microscopio óptico sobre las diferencias existentes entre las células de las criptas y las células que realizan la absorción de la superficie de la vellosidad intestinal han permitido diferenciar dos funciones: una función secretora de jugo digestivo (denominado succus entericus) y una función de absorción de sustancias nutritivas. Las células de las criptas llevarían a cabo la primera función y las células de las vellosidades la segunda. Investigaciones ulteriores han puesto de manifiesto que, si bien las células que realizan la absorción captan los productos de digestión de las sustancias nutritivas, la relación de ambos tipos de células con el líquido digestivo es más compleja de lo que en principio se supuso.

La clave de tal relación son las figuras mitóticas que se observan en las criptas. La distribución de tales figuras indica una proliferación celular poco usual en las poblaciones celulares de un organismo adulto. En 1958, Charles P. Leblond y Bernard Messier, de la Universidad McGill, demostraron que las células de las criptas migraban hacia el extremo apical de las vellosidades. Leblond inyectó en ratones timidina marcada con tritio, el isótopo radiactivo del hidrógeno. La timidina es uno de los cuatro nucleótidos que constituyen las unidades básicas de la molécula de ADN. Se incorpora en el ADN solamente cuando la célula está a punto de dividirse. De los cuatro nucleótidos del ADN la timidina es el único que no forma parte del ARN, de ahí que su captación por las células sea un signo claro de replicación del ADN. Pocas horas después de la inyección, las células marcadas con el isótopo radiactivo se encontraban únicamente en las criptas. Al cabo de 24 horas, la timidina marcada se observaba en células situadas hacia la mitad de la vellosidad y, después

de 48 horas, las células marcadas se hallaban en el extremo apical de la vellosidad intestinal.

El ciclo de proliferación y migración de las células epiteliales de la vellosidad requiere aproximadamente el mismo intervalo de tiempo en todos los mamíferos. Cuando las células alcanzan el extremo apical de la vellosidad se desprenden en la luz del tubo digestivo, acarreando con ellas los enzimas digestivos que poseen. Los enzimas que forman el jugo intestinal no se segregan, pues, de la forma habitual, sino que se vierten al lumen junto con el resto de componentes celulares. Alrededor de 17.000 millones de células se despreden diariamente a lo largo del intestino humano.

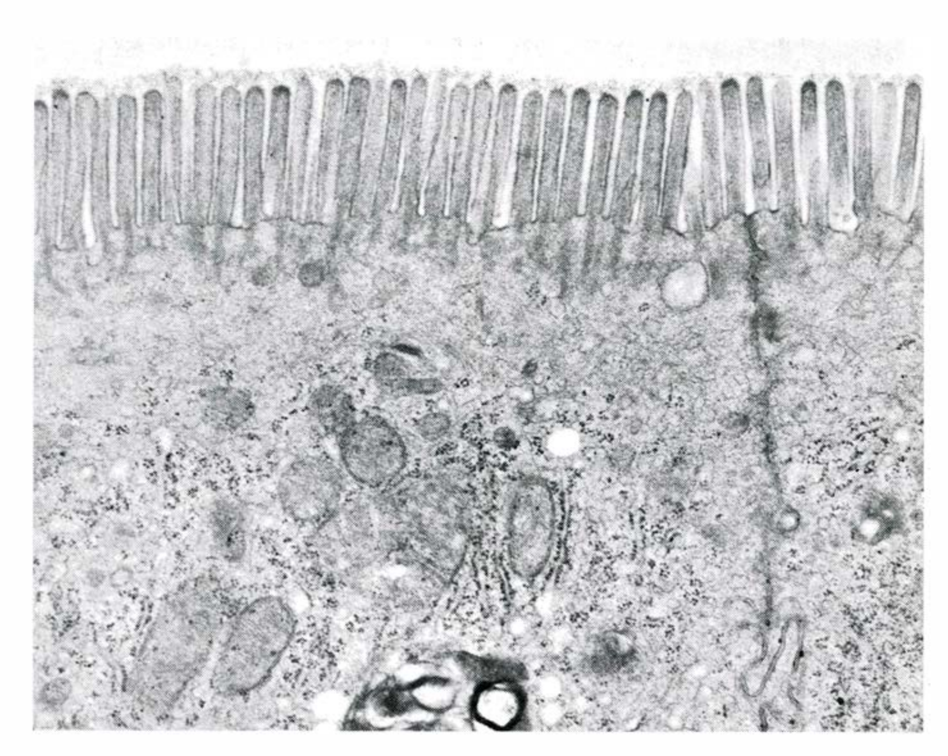
La elucidación de los mecanismos implicados en el ciclo de migración de las células epiteliales de las vellosidades intestinales ha estimulado la investigación sobre dichas células. Estudios ulte-



ESTRUCTURA DE UNA CELULA implicada en la absorción, que demuestra sus funciones especializadas en la digestión. Esta célula es uno de los dos tipos celulares que constituyen el epitelio de la vellosidad. El otro tipo es la célula caliciforme, que segrega moco. Ambos tipos se producen por división celular en las criptas y alcanzan la estructura propia de las células maduras al cabo de los dos días que dura su migración hacia la porción apical de la vellosidad. Al verse expuestas a la acción digestiva de los líquidos intestinales, parte de su membrana puede desprenderse. La célula posee abundantes mitocondrias, que proporcionan la energía requerida para los procesos metabólicos, retículo endoplasmático, donde se forman las proteínas, incluidos los enzimas digestivos, y vesículas del aparato de Golgi, donde las cadenas de carbohidratos se unen a algunas proteínas. Las células del epitelio intestinal están conectadas entre sí formando una capa única mediante bandas de cierre y desmosomas. El núcleo se encuentra en la parte profunda de la célula. La característica más sorprendente de las células es el ribete en cepillo, una estructura prominente dirigida hacia la luz intestinal desde la superficie de la célula. El ribete en cepillo está formado por una hilera de proyecciones diminutas denominadas microvilli. En la parte central de los microvilli se encuentra un haz de fibras proteicas. Las fibras de cada una de las microvellosidades se proyectan hacia abajo y se entrelazan con un sistema transversal de fibras.

riores revelaron que las células encargadas de la absorción y las células caliciformes se diferenciaban durante su migración a lo largo de la vellosidad. En el transcurso de este proceso cada tipo celular adquiere su estructura característica.

Una célula especializada en la absorción es alargada y delgada. Su núcleo está situado en las profundidades del citoplasma, por debajo de una amplia zona apical en la que abundan las mitocondrias, el retículo endoplasmático y los sacos del aparato de Golgi. La abundancia de tales orgánulos celulares indica una intensa actividad oxidativa y de síntesis. Justamente por debajo de su superficie, las células están unidas a sus vecinas por una yuxtaposición de las membranas celulares, las denominadas bandas de cierre. Las moléculas no pueden atravesar tales uniones, lo que impide que las sustancias nutritivas alcancen los vasos sanguíneos del centro de la vellosidad a través de canales intercelulares, como se creyó durante algún tiempo. Por debajo de las bandas de cierre encontramos los desmosomas, conexiones más laxas entre las membranas celulares. Ambos tipos de unión conectan las células epiteliales formando una capa única que recubre las vellosidades.



RIBETE EN CEPILLO de una célula epitelial implicada en la absorción intestinal del duodeno del ratón. La imagen, obtenida en un microscopio electrónico, está ampliada 18.000 veces. La estructura del ribete en cepillo no puede descifrarse con un microscopio óptico, ya que trasciende su poder de resolución. Hasta el año 1950 no se conoció dicha estructura. El microscopio electrónico demostró que estaba formada por proyecciones individuales denominadas microvellosidades. Tales microvilli miden aproximadamente un micrometro de longitud y se cuentan hasta 200.000 por milímetro cuadrado, en el yeyuno humano. Incrementan la superficie del intestino delgado por un factor de alrededor de 20. La superficie externa de cada microvellosidad está cubierta por el glicocálix, una capa de hidratos de carbono más gruesa que la existente en la mayoría de las células. El glicocálix aparece en la micrografía como una pelusilla que recubre la parte apical de las microvellosidades. La línea oscura de la derecha dibuja el límite que separa dos células epiteliales contiguas. La microfotografía electrónica de transmisión es del laboratorio del autor.

La característica más distintiva de las células que realizan la absorción es el borde estriado de alrededor de 1,5 micrometros de altura que recubre la superficie luminal de la célula. Su estructura está por debajo del poder de resolución del microscopio óptico e intrigó a los histólogos hasta que, en 1950, Barbara Granger y Richard F. Baker, de la Universidad de California del Sur, demostraron en el microscopio electrónico que la superficie de las células epiteliales emitía una serie de proyecciones. Estas pequeñas proyecciones, a las que se denominó microvilli (microvellosidades) medían alrededor de un micrometro de longitud y estaban empaquetadas con una densidad de alrededor de 200.000 por milímetro cuadrado en el yeyuno humano. Su presencia en el ribete en cepillo, como se llama actualmente al borde estriado de las células epiteliales, incrementa la superficie del intestino por un factor de alrededor de 20. La superficie intestinal, que se suponía de alrededor de 15 metros cuadrados, abarca en realidad unos 300 metros cuadrados.

El interior de cada microvellosidad del ribete en cepillo contiene una serie de fibras que mantienen la forma del bastoncito. Las fibras de todos los microvilli de una misma célula se prolongan hacia abajo, constituyendo una red fibrosa terminal. Cubriendo las microvellosidades se encuentra la membrana celular, formada por dos capas de moléculas lipídicas. Cada molécula posee una "cabeza" hidrofílica y dos "colas" hidrofóbicas formadas por las cadenas de ácidos grasos. En la membrana, las cabezas hidrofílicas se dirigen hacia fuera, mientras que las colas de las dos capas se encuentran frente a frente. La membrana de las microvellosidades posee un recubrimiento de hidratos de carbono que es más grueso que el propio de la mayoría de las membranas celulares.

La membrana que recubre las microvellosidades constituye la barrera real que deben atravesar las sustancias nutritivas para penetrar en el interior del organismo. Una serie de investigaciones llevadas a cabo en diversos laboratorios han puesto de manifiesto que existe todo un dispositivo de enzimas y sistemas de transporte incorporado en la membrana. Los enzimas pueden digerir una gran variedad de sustancias nutritivas, dando lugar a moléculas que los sistemas de transporte introducen en el citoplasma de las células epiteliales. A veces, las moléculas sufren una modificación ulterior en el citoplasma de la célula antes de pasar a la circulación.

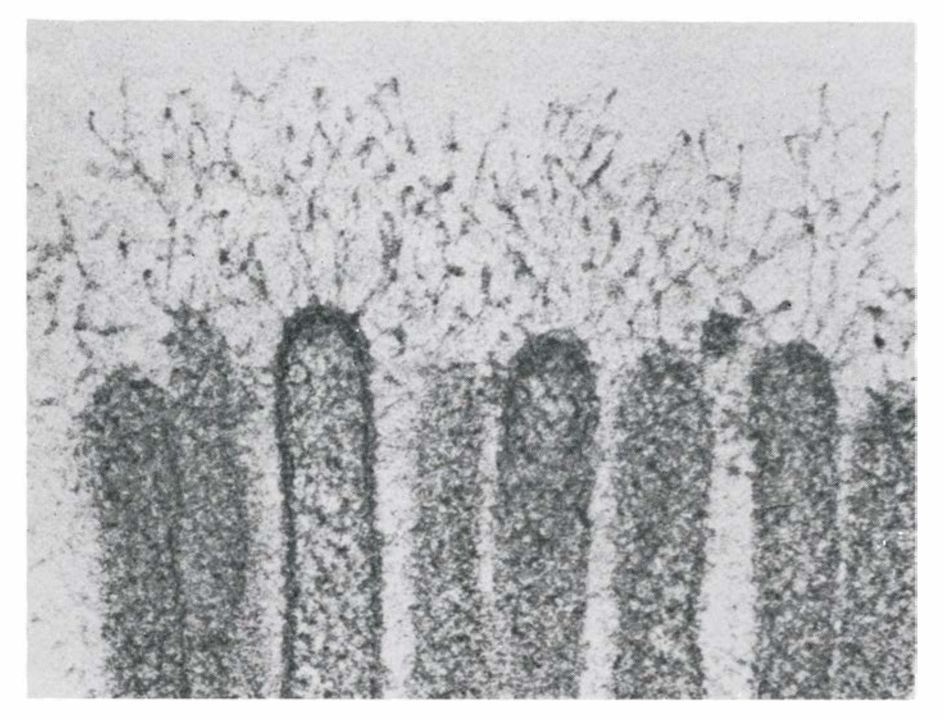
Susumu Ito, de la Universidad de Harvard, fue el primero en establecer que el recubrimiento de carbohidratos de la membrana celular conocido como glicocálix estaba compuesto de fibras muy finas que se prolongaban sin solución de continuidad a partir de la membrana celular. En la década de los cuarenta, George Gomori, de la Universidad de Chicago, y H. Takematsu, de la Universidad de Tokio, desarrollaron una técnica capaz de mostrar que un enzima podía ocupar una posición fija en la célula. Utilizaron su método para demostrar que el enzima fosfatasa alcalina, presente en las células que realizan la absorción, se localizaba en el ribete en cepillo. La fosfatasa alcalina hidroliza, es decir, rompe, una serie de compuestos fosfatados, entre ellos muchos que forman parte de la dieta humana. En los años 60 se demostró por análisis químico de los tejidos y microscopía electrónica que la fosfatasa

alcalina se encontraba únicamente en las microvellosidades de la membrana. La fosfatasa alcalina y el resto de enzimas adscritos a la membrana del ribete en cepillo son glicoproteínas, esto es, son proteínas con cadenas laterales de hidratos de carbono.

En 1961, David Miller y Robert K. Crane, de la Universidad de Washington en San Luis, idearon una técnica para aislar fragmentos intactos de ribete en cepillo a partir de homogenizados de mucosa intestinal. Aislando el ribete en cepillo, Miller y Crane podían analizarlo químicamente. Descubrieron, entre sus constituyentes, sacarasa y maltasa, ambos disacaridasas; son enzimas que rompen un azúcar de 12 átomos de carbono en moléculas de seis átomos de carbono. La sacarasa hidroliza una molécula de sacarosa, o azúcar de caña, en una molécula de glucosa y una de fructosa. La maltasa hidroliza la maltosa (un derivado del almidón) en dos moléculas de glucosa. A partir de entonces se han descubierto en el ribete en cepillo otras disacaridasas, entre ellas la lactasa, que rompe el azúcar lactosa de la leche en glucosa y galactosa. Investigaciones llevadas a cabo en otros laboratorios han informado de la presencia en el ribete en cepillo de aminopeptidasas, enzimas que liberan el aminoácido terminal de un péptido de cadena corta.

En la década de los 70 se demostró ya que los citados enzimas formaban parte integrante de la membrana de la microvellosidad, y no estaban simplemente asociados por algún tipo de interacción débil. Disponemos hoy de nuevas técnicas que nos permiten separar la membrana de las microvellosidades del resto del ribete en cepillo. Al microscopio electrónico, la membrana aislada a partir de las células que realizan la absorción aparece constituida por una serie de vesículas cerradas con numerosas protrusiones en su superficie externa. De estas preparaciones se ha obtenido una fracción que posee la actividad de los enzimas digestivos.

Las disacaridasas se liberan de las vesículas mediante tratamiento con papaína, una proteasa, es decir, un enzima que hidroliza proteínas y rompe la unión entre las disacaridasas y la membrana. La extracción de la fosfatasa alcalina y de las aminopeptidasas de las vesículas requiere la aplicación de un detergente capaz de romper las interacciones entre moléculas de lípidos. Estos hallazgos indican que las disacaridasas se encuentran unidas sólo a la parte más externa de la membrana, de ahí que una proteasa como la papaína pueda liberarlas. La fosfatasa alcalina y las aminopeptidasas estarían ancladas más profundamente en la membrana, pues se requiere, para su extracción, la ruptura de interacciones de moléculas de lípidos del interior de la membrana.

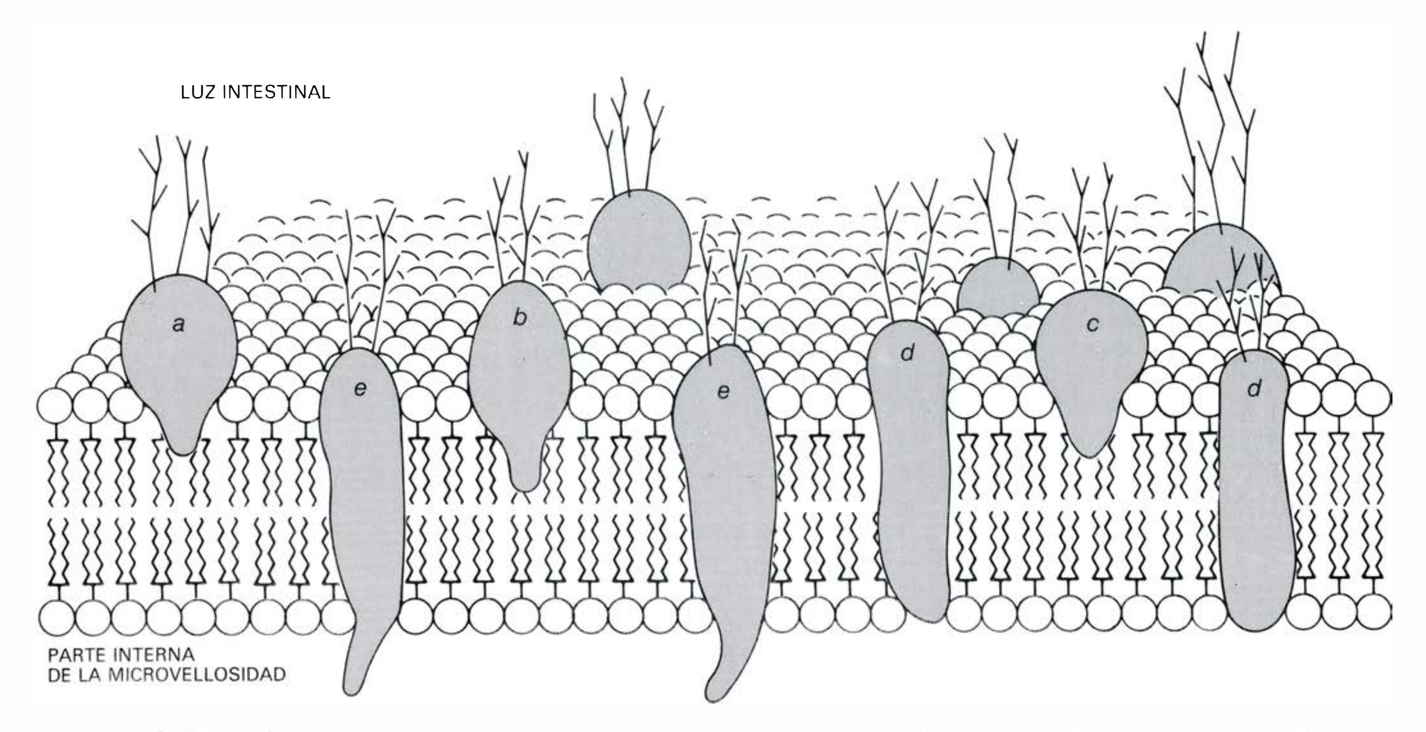


GLICOCALIX, proyectado desde la porción apical de las microvellosidades hasta la luz intestinal. La micrografía electrónica realizada en la década de 1960 por Susumu Ito, de la Universidad de Harvard, muestra las microvellosidades del íleon de un gato ampliadas 140.500 diámetros. Supuso la primera prueba de que las fibras del glicocálix estaban ancladas en la membrana de las microvellosidades. Se sabe actualmente que algunas de las fibras están constituidas por enzimas digestivos unidos a la membrana. Los enzimas son glicoproteínas, péptidos formados por una proteína con varias cadenas de carbohidratos. La porción de carbohidrato se proyecta hacia la luz intestinal. Los enzimas situados en la membrana de la microvellosidad realizan el grueso de la digestión de azúcares, proteínas y compuestos de fosfato.

S. Maroux y colaboradores, del Instituto de Bioquímica y Biología Molecular de Marsella, han ido más lejos, demostrando que las aminopeptidasas no sólo penetran en el interior de la membrana, sino que hacen protrusión en su superficie interna, proyectándose hacia la parte central de la microvellosidad. La teoría sobre las relaciones entre la membrana y los enzimas digestivos que hoy goza de mayor aceptación postula que el componente proteico de las glicoproteínas penetra en grado variable en la membrana, dependiendo del enzima, mientras que el componente carbohidratado se proyecta hacia la luz intestinal y constituye el glicocálix que aparece en las micrografías electrónicas cual una red de fibras muy finas.

Los enzimas glicoproteicos llevan a cabo la mayor parte de la función digestiva en la superficie de la microvellosidad, aunque la actividad enzimática de la membrana varíe notablemente a lo largo de la longitud de la microvellosidad. Durante algún tiempo se consideró que los enzimas de la membrana y los sistemas de transporte se sintetizaban mientras las células que realizan la absorción se encontraban aún en las criptas, y que iban entrando en funcionamiento a medida que éstas se dirigían hacia la parte apical de la vellosidad. Pero se sabe ya que la síntesis proteica continúa mientras las células epiteliales avanzan vellosidad arriba.

Milton M. Weiser desarrolló, en el laboratorio de Kurt J. Isselbacher, de Harvard, una técnica que permite recoger las distintas células que recubren la vellosidad en fracciones separadas, desde la zona apical hasta la cripta. El procedimiento consiste en llenar una porción de intestino con una solución salina tamponada, seguida de una disolución que contiene ácido etilén-diaminotetraacético (EDTA). Al cabo de una hora, las células flotan libremente, empezando por las de la región apical y terminando por las de las criptas. Por este procedimiento Weiser separó las células epiteliales en nueve fracciones y confirmó que en las células de las criptas no había actividad fosfatasa alcalina ni sacarasa, pero que aumentaba progresivamente hacia la región apical de la vellosidad.



ENZIMAS DIGESTIVOS GLICOPROTEICOS, que alcanzan distintas profundidades en la membrana de la microvellosidad. El componente proteico del enzima se inserta en la matriz lipídica que constituye la membrana; las cadenas de carbohidratos se proyectan hacia la luz intestinal. Los lípidos de la membrana están dispuestos en dos capas superpuestas, con sus "cabezas" hidrofílicas dirigidas hacia fuera y sus "colas" hidrofóbicas, cadenas de ácidos grasos, enfrentadas entre sí. Los enzimas anclados en la membrana son de varios tipos. Las distintas disacaridasas (*a, b, c*) rompen, respectivamente, un determinado azúcar de 12 carbonos en subunidades de seis. Estos enzimas se proyectan desde la membrana, donde sólo penetran hasta cierta profundidad. La fosfatasa alcalina (*d*) hidroliza muchos de los compuestos fosfatados de la ingesta; la parte proteica sí se adentra hondamente en la membrana. Las aminopeptidasas (*e*) liberan un aminoácido terminal de un péptido de cadena corta; atraviesan la membrana y alcanzan el interior de la microvellosidad.

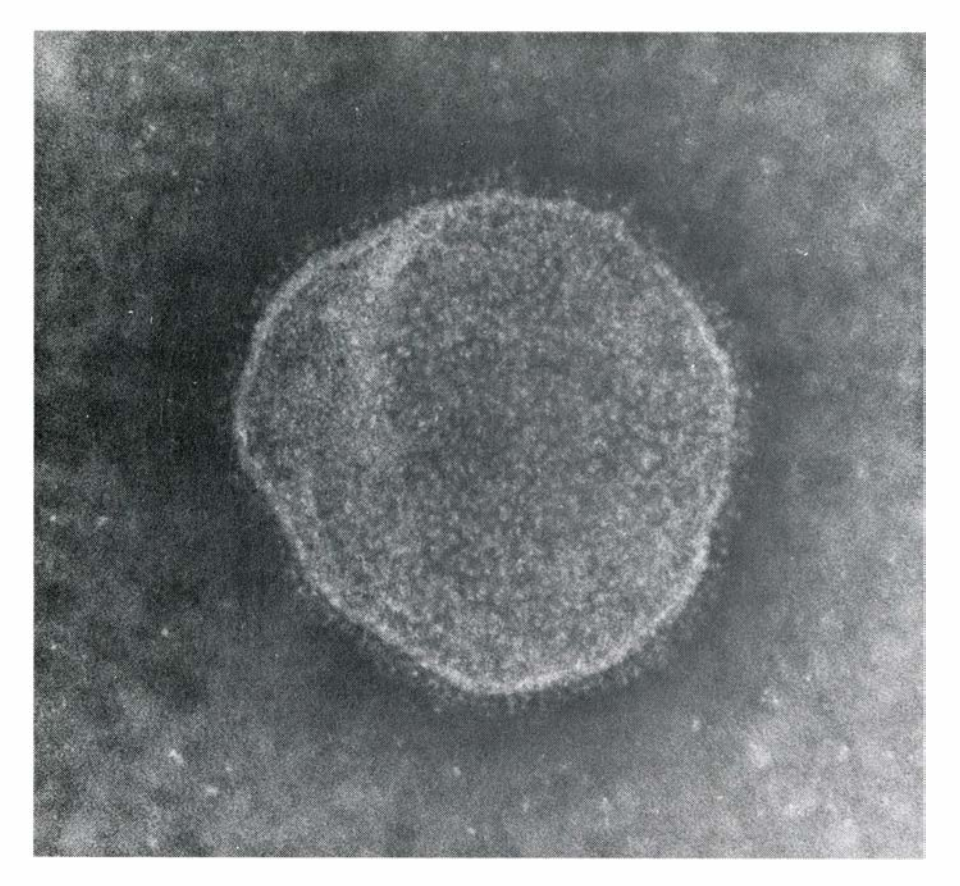
Weiser se sirvió de su técnica para investigar los procesos subyacentes a tal gradiente de actividad enzimática. Ha demostrado que la incorporación de los azúcares de seis carbonos en las glicoproteínas (que contienen fosfatasa alcalina y disacaridasas) ocurre más rápidamente en la porción apical de la vellosidad que en su región basal. Se sabía que las moléculas de azúcar marcadas eran captadas por las células que realizan la absorción y entraban en las vesículas del aparato de Golgi a los cinco minutos de haber sido inyectadas en ratas. Las vesículas del aparato de Golgi forman un sistema de membranas donde los carbohidratos se incorporan en las proteínas recién sintetizadas para constituir glicoproteínas. Finalizado el proceso, las nuevas glicoproteínas se transportan al ribete en cepillo.

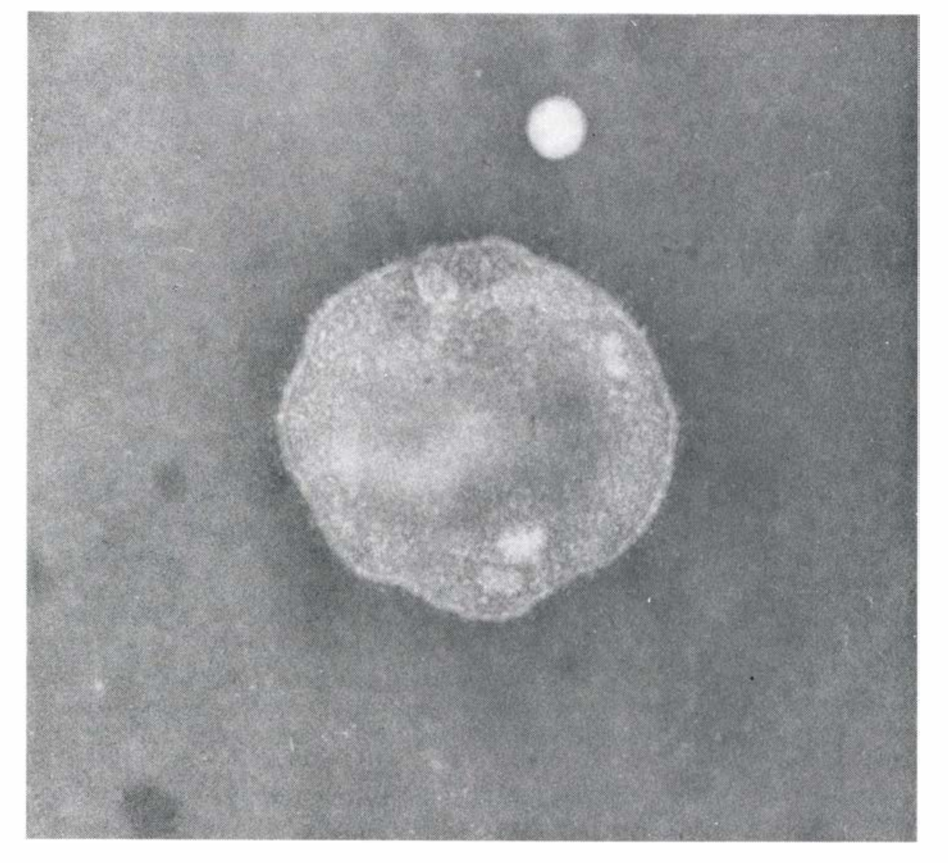
Weiser investigó la velocidad de captación del azúcar colocando las células de la vellosidad en una solución de glucosamina marcada con un isótopo radiactivo del carbono. La glucosamina es un aminoazúcar que se agrega a las cadenas laterales carbohidratadas de algunas glicoproteínas. Las células procedentes de la mitad superior de la vellosidad captaban la molécula marcada a una velocidad aproximadamente doble de la propia de las células situadas en la base. Las células de las criptas captaban una cantidad muy pequeña de azúcar marcado.

Para confirmar tales resultados, Weiser inyectó en ratas una solución de moléculas de azúcar marcadas y fue recogiendo progresivamente las distintas fracciones celulares con EDTA. Las células situadas en la región apical mostraban de nuevo una incorporación del isótopo mayor que las células basales.

El patrón de síntesis de ARN en las células de la vellosidad apoya también la teoría de una síntesis proteica más activa en la región apical. Al inyectar en ratas o pollos un isótopo radiactivo del fósforo, las células del epitelio intestinal lo captan y lo incorporan a las moléculas de ARN que se están sintetizando. Las células de las criptas asumen el fósforo radiactivo tanto en el ARN ribosómico como en el ARN mensajero. (El ARN ribosómico forma parte de la maquinaria molecular necesaria para la síntesis de todas las proteínas; el ARN mensajero especifica las secuencias de aminoácidos de las proteínas.) En la superficie de la vellosidad, sin embargo, el fósforo radiactivo se incorpora solamente en la forma del ARN mensajero denominada ARN poliadenilado (ARN-poliA). Este hecho fue descubierto en 1980 por Alan Morrison y John W. Porteus, de la Universidad de Aberdeen. Para la síntesis de cada tipo de proteína se necesita una molécula específica de ARN-poliA. Las moléculas de ARN-poliA son inestables y deben sintetizarse continuamente para garantizar la producción de una cantidad adecuada de glicoproteínas. El ARN ribosómico, por otra parte, es estable y no necesita renovarse durante el tránsito de las células desde las criptas hasta el extremo apical de la vellosidad.

¿Por qué se requiere tanta energía para sintetizar glicoproteínas en las células que realizan la absorción, que tienen una vida corta y se encuentran, aproximadamente, del todo diferenciadas? La respuesta radica en el hecho de que el ribete en cepillo, expuesto al movimiento del contenido intestinal y a los enzimas digestivos de la luz intestinal, se destruye rápidamente y debe repararse sin cesar. David H. Alpers, de la Facultad de Medicina de la Universidad de Washington, ha puesto de manifiesto que los enlaces que unen las disacaridasas a la membrana están sometidos al ataque de las proteasas segregadas por el páncreas. La vida media de una disacaridasa de la membrana es de tan sólo unas pocas horas. Además, a medida que progresa la digestión se desprenden fragmentos de la





FRACCION DE MEMBRANA de una preparación de ribete en cepillo. Aparece constituida por una serie de vesículas cerradas con múltiples proyecciones puntuales del glicocálix en su superficie externa. En la micrografía electrónica superior, las vesículas están ampliadas 263.000 veces. Si se tratan las vesículas con papaína, un enzima que rompe proteínas, las proyecciones desaparecen, tal como muestra la micrografía inferior, ampliada 237.000 veces. Pueden obtenerse disacaridasas activas a partir de las vesículas utilizando papaína, que rompe la unión que ancla los enzimas en la superficie de la membrana. En cambio, la fosfatasa alcalina que penetra a gran profundidad en la membrana, y las aminopeptidasas, que la atraviesan, sólo se liberan por medio de sustancias capaces de alterar la matriz lipídica de la membrana. Las dos micrografías fueron tomadas por Dominique Maestracci, de la Universidad de Montreal.

membrana. Microvellosidades enteras pueden romperse en fragmentos. Sólo una síntesis continua mantiene la estructura requerida para la digestión y absorción.

Cuando los enzimas de la luz intestinal y los enzimas glicoproteicos anclados en la membrana de las microvellosidades han descompuesto los alimentos en sus moléculas constituyentes, las sustancias nutritivas deben absorberse y pasar a la corriente sanguínea, desde donde alcanzarán los diferentes tejidos del organismo. Si bien se han realizado grandes progresos en el establecimiento de las relaciones existentes entre los enzimas glicoproteicos y la membrana de las microvellosidades, no se conocen bien los mecanismos por los cuales las sustancias nutritivas se transportan a través de la membrana.

El almidón, principal fuente de hidratos de carbono de la mayoría de las dietas humanas, se descompone en maltosa bajo la acción de la amilasa, tanto en la boca como en la luz intestinal. La amilasa es un enzima segregado por las glándulas salivales y el páncreas. Algunas de las moléculas de amilasa pueden quedar atrapadas entre las fibras del glicocálix junto a otros enzimas pancreáticos. Tal como se ha mencionado anteriormente, la maltasa de la membrana del ribete en cepillo rompe las moléculas de maltosa, dando lugar a dos moléculas de glucosa.

Cuando se demostró que las glicoproteínas con función enzimática formaban parte de la membrana, pareció lógico que una molécula de maltasa, tras romper el disacárido, actuara como transportador de las moléculas de seis átomos de carbono, en su movimiento hacia el interior de la microvellosidad. Tal mecanismo explicaría parte del transporte de glucosa en las células epiteliales del intestino. Se sabe, no obstante, que la maltasa de la membrana puede romper muchas más maltosas de las que el sistema de transporte podría acarrear; algunas de las moléculas de glucosa liberadas pasan, pues, a la luz intestinal.

Después de circular por la luz del intestino, la glucosa vuelve a la superficie de la vellosidad. Se supone que un segundo sistema de transporte acarrea la molécula a través de la membrana. Este segundo sistema sería capaz de bombear moléculas de glucosa al interior de las células absorbentes. El movimiento de glucosa estaría acoplado con el transporte de iones sodio en la misma dirección. Este tipo de transporte aco-

plado se ha observado también en otras células. El flujo paralelo de sodio puede proporcionar la energía necesaria para acumular glucosa en contra de un gradiente de concentración. Un defecto en este sistema adicional de transporte sería el responsable de una enfermedad denominada malabsorción de glucosa-galactosa, caracterizada por diarreas graves. Si bien se ha sugerido un esquema para este segundo mecanismo de transporte, se desconoce aún la secuencia de fenómenos moleculares que conducen a la absorción de las moléculas de azúcar.

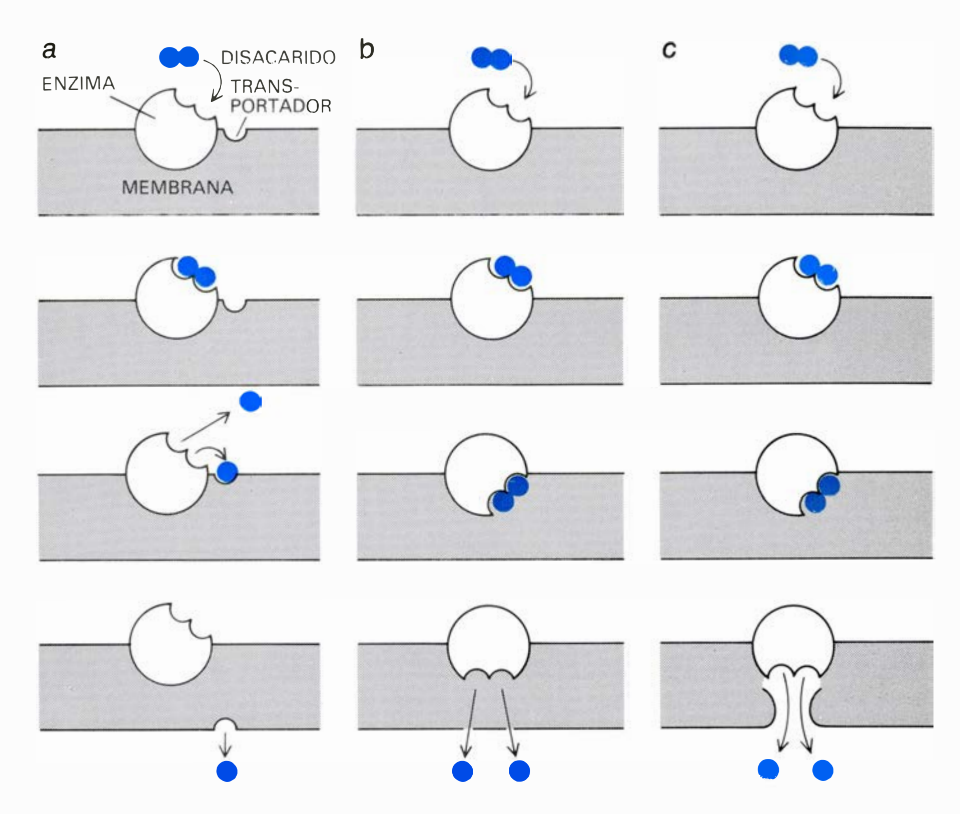
Todavía resulta más oscura la vía a través de la cual los aminoácidos y los péptidos pequeños penetran en la célula epitelial del intestino. Mientras que los carbohidratos llegan a la membrana de la microvellosidad como simples pares de azúcares de seis átomos de carbono, las proteínas se digieren parcialmente en la luz intestinal y forman una serie de distintos aminoácidos y cadenas peptídicas de pequeño tamaño. En la década de los 50 se pensaba que las proteasas segregadas en el estómago y en el páncreas digerían completamente las proteínas, reduciéndolas a sus aminoácidos componentes. La digestión en la luz intestinal da lugar en realidad a péptidos de tamaño variable y aminoácidos libres. A finales de la década de los 60, se disponía ya de datos experimentales que indicaban que las células epiteliales podían absorber directamente dipéptidos (moléculas formadas por dos aminoácidos). Ultimamente, el cuadro de la absorción de proteínas se ha complicado todavía más. Resulta ahora evidente que también los tripéptidos, formados por tres aminoácidos, pueden absorberse directamente.

La absorción de los aminoácidos parece requerir cuatro sistemas de transporte activo claramente independientes entre sí. Desconocemos la estructura molecular de algunos de ellos. Tales sistemas se corresponden con las cuatro clases distintas en que los bioquímicos clasifican la veintena de aminoácidos que se dan cita en el intestino. La clase más abundante es la de los quince aminoácidos neutros, que poseen un grupo amino ($-NH_2$) y otro carboxílico ($-COOH$) unidos al mismo átomo de carbono. Tal configuración resulta necesaria para que el transportador reconozca al aminoácido. Los aminoácidos neutros se transportan con rapidez, aunque la velocidad varía según la afinidad del aminoácido por el transportador. Las moléculas de aminoácidos neutros compiten entre sí para entrar en la célula epitelial. Los tres aminoácidos dibásicos, que poseen dos grupos amino, también pugnan entre sí por un sitio de transporte común. Pero apenas si es tal la competición cruzada entre aminoácidos neutros y dibásicos para entrar en la célula epitelial. La glicina, el aminoácido más simple, y la prolina y la hidroxiprolina, en las que el átomo de nitrógeno del grupo amino forma parte de una estructura en anillo, muestran afinidad por un tercer sistema de transporte. El ácido glutámico y el ácido aspártico, que poseen dos grupos carboxílicos, tienen también su propio sistema. Resulta difícil determinar en este último caso si los aminoácidos entran en la célula en contra de un gradiente de concentración, puesto que se convierten inmediatamente en glutamina y asparagina, no acumulándose el ácido glutámico y el ácido aspártico en el interior de la célula. El transporte de aminoácidos en los cuatro sistemas requiere energía y está acoplado a la entrada de iones sodio.

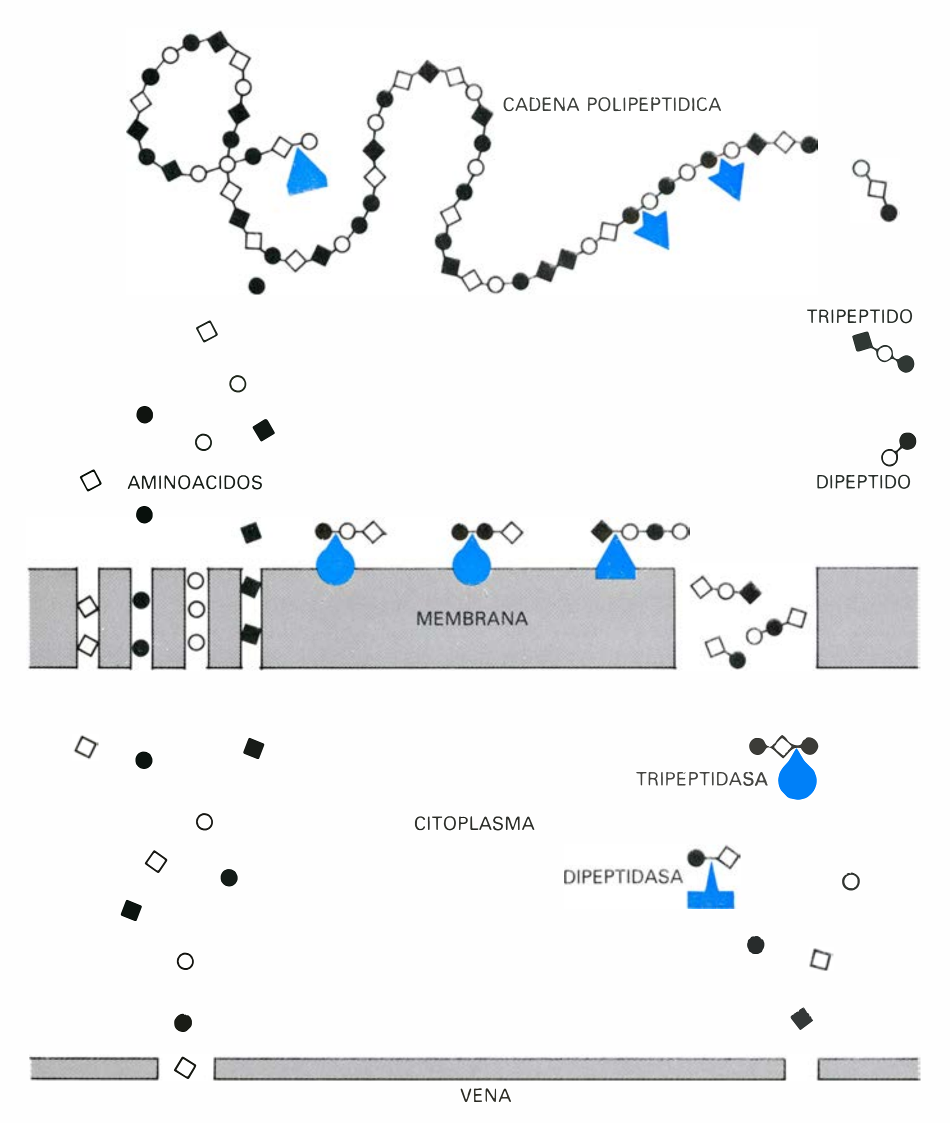
Algunas enfermedades raras prueban indirectamente la existencia de los citados sistemas de transporte, pues en ellas está afectada la absorción de una clase de aminoácido pero no la de otros. En la enfermedad de Hartnup, por ejemplo, el intestino es incapaz de absorber ciertos aminoácidos neutros, pero transporta otros sin dificultad.

Un mecanismo independiente de los que actúan en el transporte de aminoácidos acarrea dipéptidos y tripéptidos a través de la membrana de las microvellosidades. Cuando un tetrapéptido llega al ribete en cepillo, una tetrapeptidasa hidroliza su aminoácido terminal y da lugar a un tripéptido y a un aminoácido libre. Puede que el tripéptido se absorba directamente o que un segundo enzima lo hidrolice, originando un dipéptido y un aminoácido libre. Los enzimas que hidrolizan tripéptidos están distribuidos por igual entre la membrana y el citoplasma. Los enzimas que hidrolizan dipéptidos en dos aminoácidos se encuentran predominantemente en el interior de la célula. Pocos son los dipéptidos que se hidrolizan en la misma membrana; la mayoría se transporta hacia el interior de la célula.

El transporte de péptidos en lugar de aminoácidos ofrece varias ventajas al



SE HAN PROPUESTO VARIOS MECANISMOS para explicar cómo se transportan las moléculas de azúcar a través de la membrana de las microvellosidades. La disacaridasa, que descompone azúcares de 12 átomos de carbono, como la sacarosa, en moléculas de seis átomos de carbono, está unida superficialmente a la membrana. Si la orientación del enzima permanece invariable durante el proceso de hidrólisis, los dos azúcares de seis átomos de carbono se liberarán inevitablemente en el espacio existente entre las microvellosidades (*a*). Un mecanismo independiente del enzima glicoproteico los transportaría luego a través de la membrana. Otra posibilidad consiste en que el sitio activo del enzima con el que se combina la molécula de azúcar experimente una rotación hacia el interior de la célula mientras se descompone la molécula de azúcar. En tal caso, los azúcares de 6 átomos de carbono se liberarían en la propia membrana (*b*) o en un canal que atravesara la membrana hasta el citoplasma de la célula absorbente (*c*).



AMINOACIDOS Y PEPTIDOS DE CADENA CORTA, productos de la digestión de proteínas, atraviesan la membrana de la microvellosidad por medio de distintos sistemas de transporte. Los enzimas proteolíticos de la luz intestinal hidrolizan las proteínas dando lugar a aminoácidos libres y cadenas peptídicas. La membrana de las microvellosidades posee cuatro sistemas para transportar los aminoácidos libres hasta el citoplasma de la célula del epitelio intestinal. Todos los aminoácidos con estructura química similar compiten por el mismo sistema de transporte. El movimiento de aminoácidos hacia el interior de la célula está acoplado con un flujo paralelo de iones sodio, un proceso que proporciona la energía necesaria para acumular aminoácidos en contra de un gradiente de concentración. Las aminopeptidasas ancladas en la membrana hidrolizan el aminoácido terminal de los péptidos de cadena corta. Cuando la cadena ha experimentado un acortamiento puede pasar de nuevo a la luz o atravesar directamente la membrana. Existe un mecanismo independiente para el transporte de dipéptidos y tripéptidos. Una vez atravesada la membrana, los péptidos pequeños se descomponen en aminoácidos por la acción de las peptidasas del citoplasma. La estructura molecular detallada de los transportadores permanece todavía obscura.

organismo. El carácter competitivo del transporte de aminoácidos se evita al pasar péptidos intactos a través de la membrana. La presencia de transportadores de péptidos minimiza el gasto energético del transporte contra gradiente; los dipéptidos y tripéptidos se hidrolizan rápidamente en el interior de la célula, por lo que la concentración de péptidos en el citoplasma no sobrepasa la existente en la luz intestinal. Además, cualquier deficiencia en el sistema de transporte de los aminoácidos aislados puede quedar compensada por el transporte de péptidos, puesto que las moléculas que no pueden entrar por una alteración del sistema de transporte lo consiguen en forma de péptidos. La estructura molecular de los sistemas de transporte de péptidos aún no se ha dilucidado.

También existen lagunas en nuestro conocimiento de la estructura interna de las microvellosidades. Hasta los primeros años de la década de los setenta se pensaba que las microvellosidades eran estructuras rígidas cuya única función consistía en ampliar la superficie de la membrana. En 1975, Mark Mooseker, a la sazón en la Universidad de Pennsylvania, demostró que una preparación aislada de ribete en cepillo podía contraerse en determinadas circunstancias. Las condiciones necesarias incluían la presencia de iones calcio y de adenosín trifosfato (ATP), la molécula que interviene en las reacciones del metabolismo celular que requieren aporte energético.

Cuando las microvellosidades se contraen, las fibras de su parte central se deslizan en dirección a la red terminal subyacente. Mooseker y Lewis G. Tilney, también de la Universidad de Pennsylvania, demostraron que el ribete en cepillo contenía actina y miosina, las proteínas que se deslizan en la contracción del músculo. En el tejido muscular, una tercera proteína, la meromiosina, media en la interacción entre la actina y la miosina. Mooseker y Tilney aplicaron meromiosina a la microvellosidad para determinar si la interacción era similar a la observada en el músculo. Descubrieron que así era, en efecto, y propusieron que en la contracción de la microvellosidad los filamentos de actina de la parte central se dirigirían hacia los haces horizontales de miosina del retículo terminal. Postularon que la meromiosina facilitaría la contracción de la microvellosidad permitiendo a la actina deslizarse a lo largo de las fibras de miosina.

En el transcurso de las investigaciones encaminadas a explicar cómo se contrae la microvellosidad, se han aislado otras cuatro proteínas. De todas ellas, la más intrigante es la calmodulina, una proteína similar a la troponina, que muestra afinidad por los iones de calcio y que también regula la contracción del músculo esquelético. La calmodulina desempeña funciones análogas a la troponina en células no musculares. Quizás intervenga la calmodulina en la regulación de la contracción de las microvellosidades, aunque se ha adelantado otra explicación de su presencia. La calmodulina se concentra en el eje de la microvellosidad y no en la red terminal subyacente, donde interacciona la actina con la miosina. Puesto que el calcio es una de las sustancias que debe absorberse a través del epitelio intestinal, Mooseker ha sugerido que la calmodulina formaría parte de los sistemas de transporte de los iones calcio. Transportando dichos iones, la proteína migraría desde la parte central de la microvellosidad hasta la red terminal. Si esta hipótesis se demuestra cierta, supondría un avance en el conocimiento de los sistemas de transporte de la membrana y del citoplasma de las células del epitelio intestinal.

Temas metamágicos

Los atractores extraños son configuraciones matemáticas en equilibrio entre orden y caos

Douglas R. Hofstadter

Hace algunos meses, recorriendo en compañía de un amigo los pasillos del departamento de física de la Universidad de Chicago, me saltó a la vista el cartel anunciador de un simposio internacional sobre "Atractores extraños". Tan insólito título no pudo dejar de atraerme y extrañarme, y por ello inquirí de mi amigo mayores detalles. Según me explicó, se trataba de un tema candente en la física teórica de nuestros días. A juzgar por sus explicaciones, el tema prometía ser tan atractivo como misterioso.

Saqué la impresión de que la noción clave residía en los que pudiéramos llamar bucles de realimentación matemática: expresiones cuyas "salidas" pueden reinyectarse como nuevas "entradas", de igual forma que los sonidos emitidos por un altavoz pueden captarse por un micrófono, reprocesarse y volverse a emitir. Según parecía, incluso el más sencillo de tales bucles podía generar tanto pautas regulares y estables como organizaciones caóticas (¡admitiendo que la idea de organización caótica no sea auto-contradictoria!). Que sucediera de una u otra forma dependía del valor de un único parámetro. Pequeñas variaciones de tal valor podían provocar las disparidades más extremas en el comportamiento del sistema realimentado. La idea de ver al orden derretirse y fundirse en caos, de ver la regularidad progresivamente disolverse en la aleatoriedad, me resultaba apasionante.

Más todavía. Al parecer, acababan de salir a la luz ciertos rasgos "universales" e inesperados, típicos de la transición del orden al caos; rasgos que tan sólo dependían de la existencia de retroalimentación y virtualmente insensibles a otros aspectos del sistema. Tan gran generalidad es de la mayor importancia, porque disponiendo de modelos matemáticos capaces de mostrar gradualmente la aparición de comportamientos caóticos tal vez podamos percibir las nociones clave para comprender cómo y por qué se originan turbulencias en todo tipo de sistemas físicos. En oposición a casi todos los fenómenos que hasta ahora comprendemos mejor y más profundamente, las turbulencias son fenómenos no-lineales. Lo que significa que al sumar dos soluciones de las ecuaciones que las rigen no es forzoso que resulte una nueva solución. La comprensión de la matemática no lineal va muy por detrás de las teorías lineales; por esta causa, los físicos han venido careciendo de buenas descripciones matemáticas de la turbulencia. De lograrlas, su importancia sería fundamental.

Más tarde, cuando empecé a leer sobre estos temas, descubrí que en realidad se habían desarrollado simultáneamente a partir de muchas disciplinas. En matemática pura se habían empezado a estudiar procesos de iteración en sistemas no lineales, con auxilio de ordenadores. En meteorología teórica y en genética de poblaciones, por no mencionar campos de la física tan distantes como las teorías de fluidos, los láseres y las órbitas planetarias, se había llegado –cada cual por su propio sendero– hasta modelos matemáticos semejantes, todos ellos no lineales, caracterizados por contener bucles de realimentación, y con ellos, la semilla del caos. Cada grupo, independientemente, había estudiado sus propiedades y descubierto aspectos singulares que habían escapado a la atención de los demás. Por otra parte, no sólo los teóricos, sino también los experimentalistas de estas disciplinas, en principio tan desconectadas, habían observado fenómenos caóticos que compartían ciertas pautas de organización básica. Pronto pude ver que la simplicidad de las ideas subyacentes les confería una elegancia que, a mi juicio, rivalizaba con la de algunas de las nociones más logradas de la matemática clásica. Y, en efecto, ciertas partes de este trabajo exhalan un aroma clásico, refrescante y concreto, que será bienvenido en nuestros días, tan agobiados de abstracciones.

Es probable que la causa principal de que se exploren estas ideas precisamente ahora sea la modernidad misma del método de su descubrimiento. Es una especie de matemática experimental, donde el ordenador digital hace el papel de la nave de Magallanes, del telescopio del astrónomo, o de los aceleradores de partículas de los físicos. Al igual que las naves, telescopios y aceleradores han tenido que ir haciéndose constantemente mayores, más potentes y costosos, para poder calar en regiones cada vez más recónditas de la naturaleza, también irán haciendo falta ordenadores de tamaño, rapidez y precisión siempre crecientes para explorar regiones más y más alejadas en el espacio matemático. Por lo mismo, al igual que hubo una edad de oro de la náutica en la exploración terrestre, de la telescopía en astronomía, o de los aceleradores en la física de partículas –edad caracterizada por un pico en la gráfica de descubrimientos nuevos en función de la inversión requerida–, también podríamos esperar que con el estudio de los modelos de caos se produjera una edad de oro de la matemática experimental. Puede que tal edad haya pasado ya, o quizá, que esté en pleno desarrollo en este momento. Y tal vez seamos testigos después de una repentina fiebre de trabajos teóricos, encaminados a fundamentar y organizar estos descubrimientos experimentales.

Sea como fuere, se está desarrollando una rama de la matemática tan curiosa como amena. Esta forma de creación matemática aporta a la inteligencia una variada imaginería y poderosas intuiciones. Al poder cortocircuitar el estilo matemático tradicional –teorema, demostración, teorema, demostración– gracias a la potencia de cálculo de los ordenadores, es posible llegar rápidamente a observaciones y descubrimientos que, reforzándose mutuamente, pueden formar un entramado de ideas tan rico y coherente que, a la larga, resulte más fácil atinar con las necesarias demostraciones, pues se dispondrá ya de un "atlas" minucioso del territorio conceptual analizado. Uno de los más decididos partidarios de esta forma de matematización ha sido Stanislaw M. Ulam, quien, siendo los ordenadores todavía adolescentes, los lanzó sobre problemas de iteración no lineales, además de trastear con ellos en otras muchas regiones de la matemática. Muchas de las ideas esbozadas aquí son consecuencia de los trabajos de roturación que Ulam y Paul Stein realizaron conjuntamente.

Y basta ya de historias edificantes.

Para remontarnos hasta el concepto de atractor extraño conviene examinar antes la noción de atractor. Toda la teoría de atractores se inspira en la iteración de funciones de variable real, vale decir, en el estudio del comportamiento de sucesiones x, $f(x)$, $f(f(x))$, $f(f(f(x)))$, ..., siendo f alguna función interesante. El valor inicial de x será llamado "semilla". La idea consiste en reinyectar la salida de f, utilizándola una y otra vez como nuevo valor de entrada, con la intención de observar si se presenta alguna regularidad.

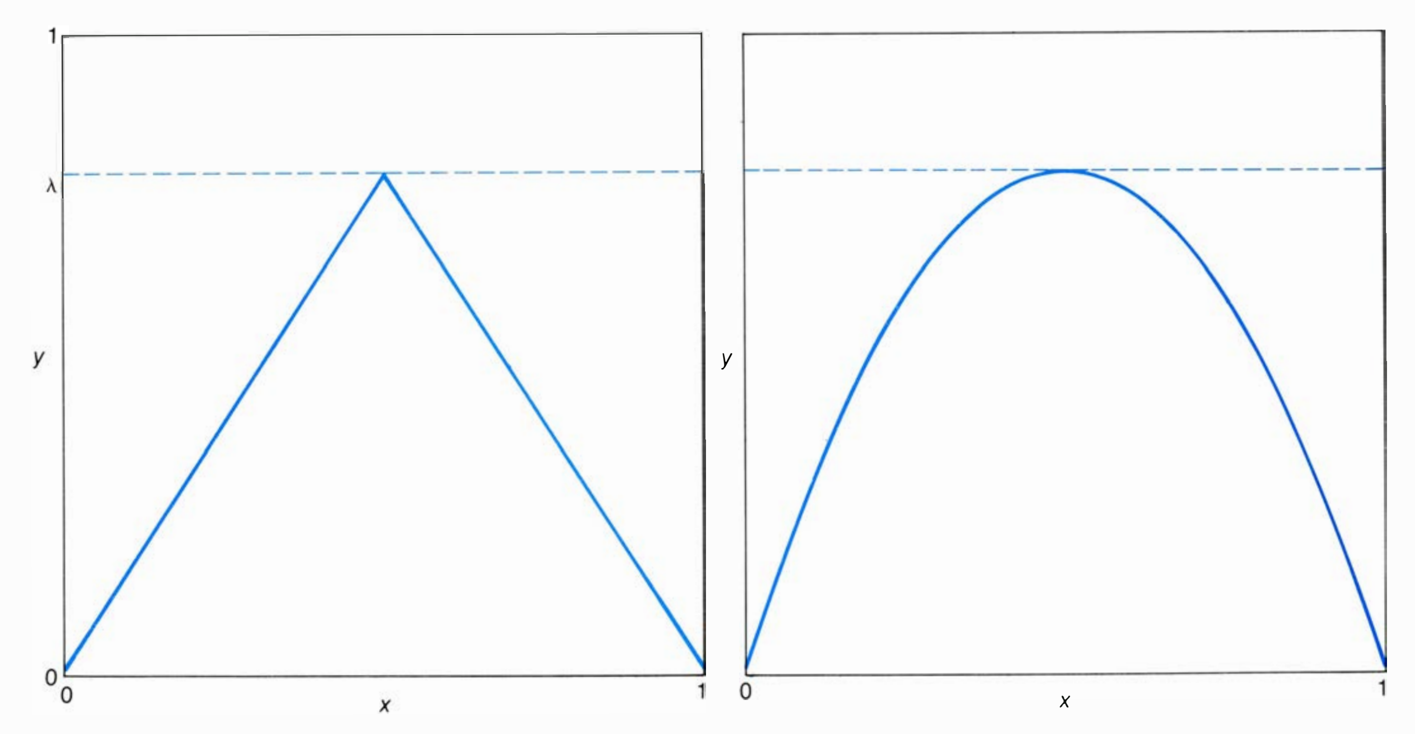
He aquí un problema interesante y no demasiado difícil concerniente a la iteración de funciones. ¿Se podrá inventar una función p con la propiedad de que para todo valor de x, $p(x)$ sea también un valor real, y se cumpla $p(p(x)) = -x$? Si el problema no es trivial, ello se debe justamente a la condición de que $p(x)$ haya de tomar valores reales, pues si no, bastaría que $p(x)$ fuera ix (siendo i raíz cuadrada de -1) para tener una solución sencilla. De hecho, podemos pensar incluso que el problema consiste en buscar alguna "raíz del signo menos". Otro problema afín consiste en hallar una función real q tal que $q(q(x) = 1/x$ para todo valor de x distinto de cero. Observemos que cualesquiera que sean las funciones solución p y q que se den, su iteración genera ciclos de longitud cuatro.

Con mayor generalidad, ¿cuáles son los tipos de funciones para los que podemos esperar se manifiesten comportamientos cíclicos o cuasi-cíclicos al someterlas a iteración indefinidamente? Las funciones más sencillas, como $3x$ o x^3, no se comportan así. La n-ésima iteración de $3x$, por ejemplo, nos dará $3 \times 3 \times 3... \times 3 \times x$, donde el coeficiente repite n veces la cifra 3, es decir, $3^n x$; la n-ésima iteración de x^3 es sencillamente $(((x^3)^3)^3...)^3$, o sea, x^{3^n}, pues el exponente 3 figura n veces consecutivas. No vemos aquí rasgos cíclicos. Por ejemplo, en la primera función, si la semilla es positiva, los valores iterados crecen indefinidamente. Para invertir esta tendencia al ascenso, haría falta que la función mostrase alguna clase de vuelta atrás, que su gráfica refleje algún zigzag u oscilación. Con terminología más técnica: las funciones a utilizar *no* deben ser *monótonas*; las gráficas de nuestras funciones deben presentar gibas o pliegues; pueden comenzar, por ejemplo, ascendiendo y más tarde curvarse hacia bajo y descender.

En la figura de la parte izquierda de esta misma página tenemos un diente de sierra, cuya cúspide es una afilada punta, y en la figura de la derecha, un arco de parábola, suavemente curvado. Cada una de estas líneas parte del origen, llega a alcanzar una altura máxima, que llamamos λ, y después desciende, aterrizando en el otro extremo del intervalo. Es evidente que existe una infinidad no numerable de curvas de distintas formas que, tras alcanzar una altura máxima λ, descienden nuevamente a 0, pero seguramente estas dos se encuentren entre las más sencillas. Y de ambas, quizá sea la parábola la más sugestiva desde el punto de vista matemático. Su ecuación es $y = 4\lambda x(1 - x)$, donde λ representa un parámetro comprendido entre 0 y 1.

Tan sólo admitiremos entradas (es decir, valores de x) comprendidos entre 0 y 1. Como muestra el gráfico, para todo x contenido en ese intervalo, la salida y se encuentra siempre entre 0 y 1. Por consiguiente, el valor de salida puede ser reinyectado como valor de entrada en la función, asegurando de esta forma la posibilidad de iteración indefinida. Al iterar repetidamente funciones "jorobadas" como ésta, los sucesivos valores de y que van obteniéndose crecen unas veces, y otras decrecen, oscilando siempre entre 0 y λ. Como veremos, es precisamente la comba de la gráfica la causante de los curiosos efectos de la iteración.

Resulta que las espectaculares diferencias en el grado de regularidad de las configuraciones iterativas ya mencionadas se deben a variaciones en el ajuste del que podríamos llamar "control λ". De acuerdo con los valores en que se sitúe este mando, la función genera una increíble variedad de "órbitas", es decir, de sucesiones x, $f(x)$, $f(f(x))$, etcétera. En particular, cuando λ es inferior a un cierto valor crítico ($\lambda_c = 0{,}892486417967...$) todas las órbitas trazan pautas regulares (aunque el grado de regularidad sea muy variable; por lo común, cuanto menor es λ tanto más sencilla es la regularidad de la órbita), pero cuando λ alcanza o rebasa este valor crítico, ¡agárrense al sombrero!

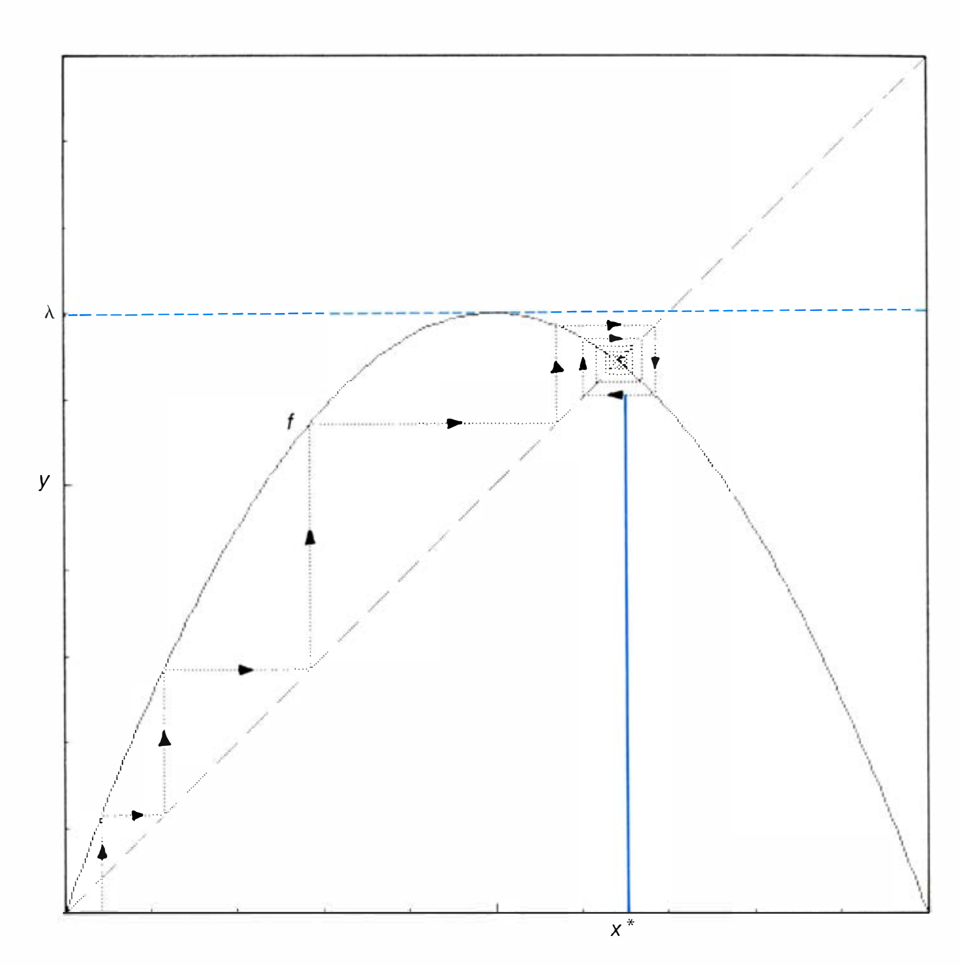


*Entre las funciones "combadas" las hay de gráfica angulosa (***izquierda***), o suave, como la parábola (***derecha***)*

La sucesión de valores x, $f(x)$, $f(f(x))$,... irá oscilando de forma esencialmente caótica, independientemente de la semilla x (de valor entre 0 y 1) que decidamos sembrar. En el caso de las parábolas, parece ser que el primero en comprender el papel crítico del mando de control λ fue P. J. Myrberg, a principios de los años 60, pero su trabajo apareció en una oscura revista y pasó desapercibido. Unos 10 años después, Nicholas C. Metropolis, Paul Stein y Myron Stein redescubrieron el importante papel desempeñado por λ, no sólo en las parábolas, sino también en muchas otras funciones. En realidad, la esencia de su descubrimiento es que en tanto se cumplan ciertas condiciones topológicas, la función particular carece de interés –tan sólo son importantes los valores de λ. Este hecho ha sido bautizado como "propiedad de universalidad estructural".

Para comprender esa dependencia tan poco intuitiva del comportamiento de la sucesión de salida respecto del valor del parámetro, es preciso formarse una imagen visual del proceso de iteración de $f(x)$. Ello se logra fácilmente. Supongamos ajustado el control de λ a 0,7. Podemos ver la gráfica de $f(x)$ en la ilustración de la derecha. Vemos también en línea de trazos la recta $y = x$, bisectriz del primer cuadrante. (Esta gráfica, así como la mayoría de las reproducidas en el artículo, han sido obtenidas en un pequeño ordenador por Mitchell J. Feigenbaum, de Los Alamos National Laboratory.)

Fijémonos en los dos valores de x donde se cortan la recta y el arco de parábola. Se encuentren en $x = 0$ y también en $x = 9/14 \simeq 0{,}643$. Denotemos por x^* al valor no nulo. Por construcción, $f(x^*) = x^*$; la iteración indefinida de f en este valor de x nos hace entrar en un bucle infinito. Lo mismo sucede al empezar la iteración en $x = 0$. Empero, hay una importante diferencia entre estos dos "puntos fijos" de f. Podemos verlo más claramente tomando algún otro valor inicial de x, por ejemplo, algún valor próximo a 0,04, como muestra la ilustración. Llamemos x_0 a este valor inicial de x. Podemos engendrar la órbita de una semilla x_0 cualquiera mediante una elegante construcción geométrica. Una recta vertical trazada por la abscisa de valor x_0 alcanzará la curva a la altura $y_0 = f(x_0)$. Para iterar $f(x)$ debemos trazar una nueva recta vertical que pase por un valor de x igual al valor y_0 de y. Aquí es donde la bisectriz $y = x$ del primer cuadrante muestra su utilidad. Situados como estamos a la altura y_0 nos desplazamos horizontalmente hasta interceptar la recta. Puesto que, a lo largo de ella, y es igual a x, tanto x como y serán en el punto de intersección iguales a y_0. Llamemos x_1 al nuevo valor de x. Tracemos ahora una segunda vertical que pase por x_1; esta recta cortará a la curva a la altura en que $y_1 = f(x_1) = f(y_0) = f(f(x_0))$. Repetimos luego el proceso.



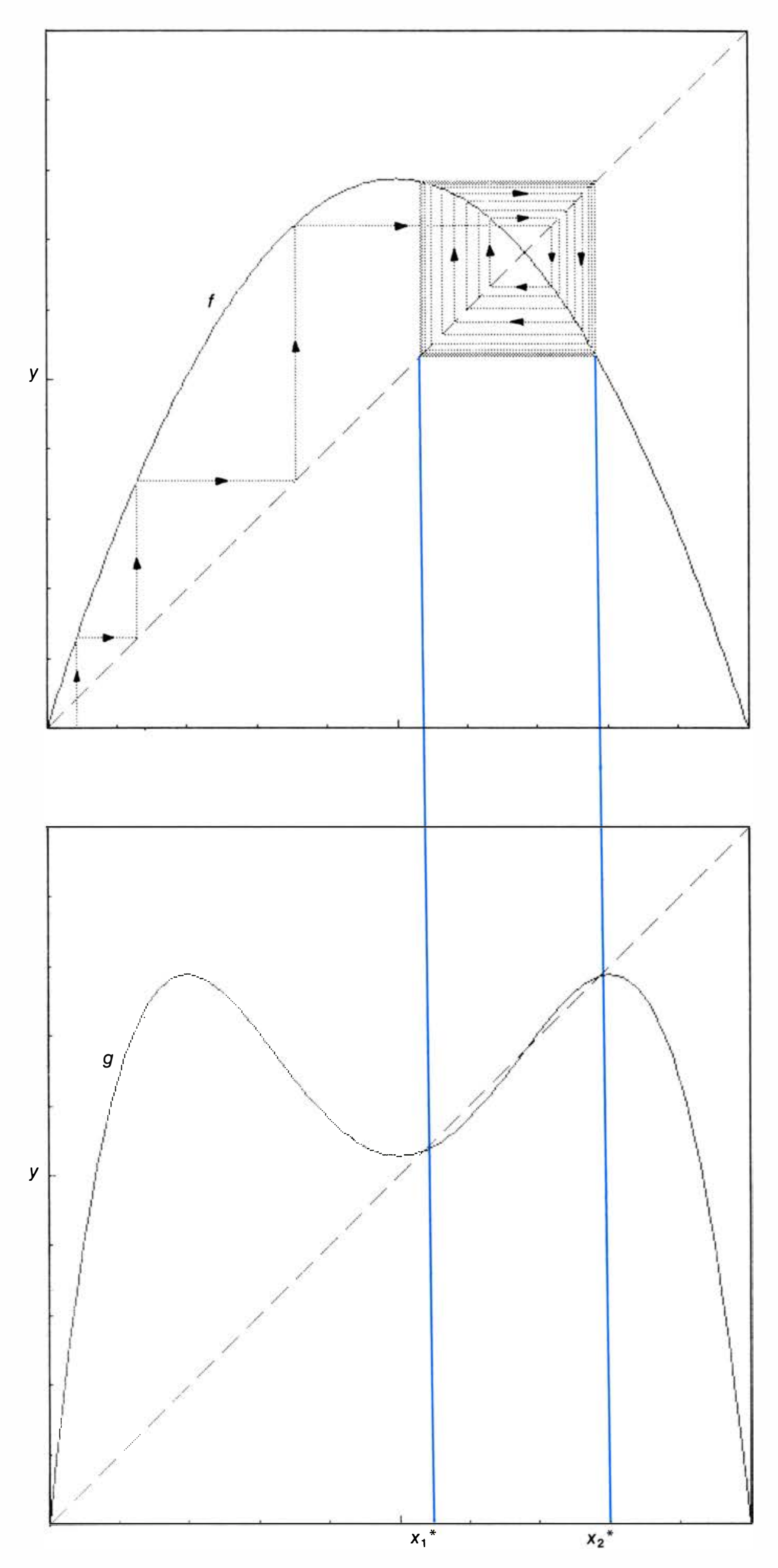
Situando el "control λ" en 0,7, la función **f(x)** *genera esta gráfica*

En breve, la receta para realizar gráficamente el proceso de iteración es sencilla: (1) Desplazamiento vertical hasta alcanzar la curva, seguido de (2) desplazamiento horizontal hasta la diagonal. Repetición, una y otra vez, de los pasos (1) y (2).

En la figura podemos ver el resultado del proceso al sembrar en él la semilla $x_0 = 0{,}04$. Nos vemos llevados una y otra vez, como en un tiovivo, alrededor del punto cuyas coordenadas x e y son ambas iguales a x^*, acercándonos gradualmente a él. Por ello, x^* es un punto fijo de tipo especial, pues atrae a sí los valores iterados de $f(x)$. Es un *atractor* del tipo más sencillo: toda semilla (exceptuada $x = 0$) se ve arrastrada por iteración de f hacia dicho valor estable de x. Consecuentemente, el valor x^* recibe el nombre de punto fijo estable. En cambio, el otro punto fijo, 0, es repelente, o inestable, pues la órbita de cualquier valor inicial de x, incluso los infinitesimalmente distintos de 0, procederán a *alejarse* de 0 y dirigirse hacia x^*. Observemos que los valores de f unas veces rebasan x^* y otras caen cortos, pero van inexorablemente aproximándose a x^*.

¿Cuál es la causa de tan radical diferencia cualitativa entre los dos puntos fijos (0 y x^*) de f? Un vistazo a la ilustración nos hace comprender que en 0 la curva está demasiado empinada. En concreto, en ese punto la inclinación es mayor de 45 grados; y es precisamente la inclinación local de la curva lo que controla la amplitud del desplazamiento horizontal en cada iteración de f. Siempre que la pendiente de la curva exceda de 45 grados (tanto si el tramo es ascendente como si es descendente) cada iteración tenderá a irnos alejando más y más del punto de partida. Por tanto, el criterio de estabilidad señala: para que un punto fijo sea estable, la inclinación de la curva en él ha de ser menor de 45 grados. Esto precisamente es lo que sucede cuando λ toma el valor 0,7: en el punto x^* la pendiente es de unos 41 grados, mientras en $x = 0$ es mucho mayor de 45 grados.

¿Qué sucederá al aumentar λ? El emplazamiento de x^* (que por defini-



Arriba: una espiral convergente a un 2-ciclo estable. Abajo vemos los elementos **x^*_1** *y* **x^*_2** *que definen el ciclo*

ción es la abscisa del punto donde se cortan la curva f y la bisectriz $y = x$ del primer cuadrante) cambiará y, al mismo tiempo, aumentará también la *pendiente* de f en el punto x^*. Cuando λ valga 3/4, esta pendiente alcanzará el ángulo de 45 grados. ¿Qué ocurrirá entonces? Llamemos Λ_1 a este valor particular del parámetro λ. Para hacernos idea de lo que podemos esperar echemos un vistazo a una figura construida para un valor ligeramente mayor, $\lambda = 0{,}785$ [*véase la ilustración de la izquierda*].

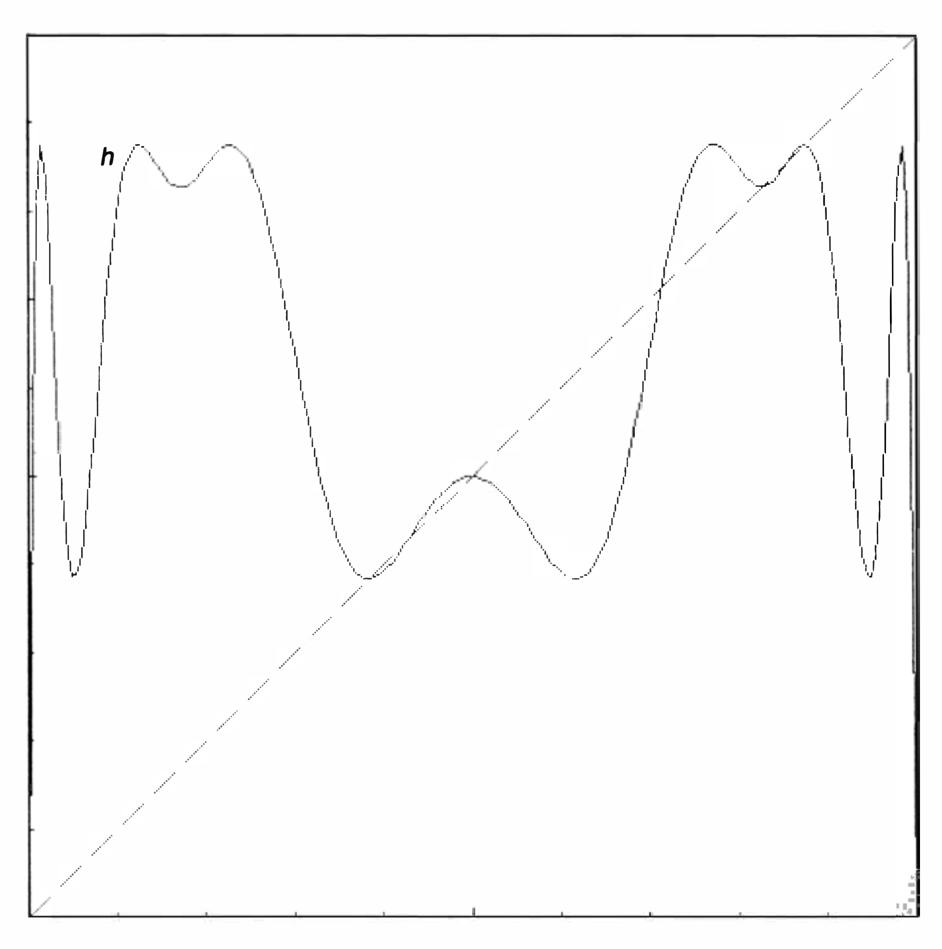
Tomemos al azar una semilla, por ejemplo, 0,04, como antes. En la mitad superior de la ilustración podemos ver la órbita resultante. Observamos en ella algo muy curioso. Al principio los valores de x van dirigiéndose a la vecindad de x^* (que sigue siendo punto fijo de f, aunque ahora inestable), pero después el proceso de iteración va describiendo una espiral cada vez algo más amplia, que tiende suavemente a estabilizarse en una especie de baile cuadrangular convergente hacia dos valores especiales, x^*_1 y x^*_2. Esta oscilación tan elegante se llama un 2-ciclo, y el par de valores que lo definen (x^*_1 y x^*_2) siguen formando un sistema atractor (o más brevemente, un atractor), que ahora tiene período dos. Con este término queremos sugerir que nuestro 2-ciclo es estable: es capaz de atraer hacia sí valores muy alejados de x por iteración de f. La órbita de toda semilla positiva (exceptuado el punto fijo x^*) terminará ejecutando esta misma danza, es decir, se aproximará asintóticamente al 2-ciclo perfecto compuesto por los puntos x^*_1 y x^*_2, si bien nunca lo alcanzará exactamente. Empero, desde un enfoque físico, la aproximación llega pronto a ser tan grande que podemos admitir sin inconveniente que las órbitas han quedado "atrapadas" por el atractor.

Para mejor comprender lo que sucede, resultará muy esclarecedor echar un vistazo a la gráfica de una nueva función generada a partir de la dada, a saber, $g(x) = f(f(x))$, que podemos ver en la parte inferior de la ilustración. Este camello de dos jorobas es la *iterada* de f. Ante todo, fijémonos en que todo punto fijo de f es también punto fijo de g; por tanto, 0 y x^* serán puntos fijos de g. Pero, en segundo lugar, observemos que como $f(x^*_1)$ es igual a x^*_2, y que $f(x^*_2)$ es igual a x_1, resulta que g tiene dos nuevos puntos fijos: $g(x^*_1) = x^*_1$ y $g(x^*_2) = x^*_2$. Hallar gráficamente x^*_1 y x^*_2 es cosa fácil: son las abscisas de los puntos donde la bisectriz intercepta a la gráfica de $g(x)$.

Hay cuatro puntos de intersección; los otros dos son 0 y x^*. Como hemos visto, el criterio de estabilidad frente a iteración de los puntos fijos es que la pendiente de la curva en estos puntos es menor de 45 grados. Ahora estamos interesados en los puntos fijos de g y, por consiguiente, en la pendiente de g y no en la de f. En la misma ilustración vemos claramente que en 0 y en x^* la pendiente de g es mayor de 45 grados, mientras que tanto en x^*_1 como en x^*_2 es menor. Más todavía, no sólo es la inclinación en ambos puntos menor que 45 grados, sino que, y vale la pena destacarlo, las pendientes son iguales ("esclavas" una de otra, se dice a veces), como es fácil comprobar calculando la derivada en estos puntos.

Acabamos de ver que un atractor de período 1 puede convertirse en un atractor de período 2 al trasponer λ un cierto valor umbral ($\lambda = 3/4$). Precisamente para este valor el punto x^* se escinde en dos valores oscilantes, x^*_1 y x^*_2. Es obvio que en su "nacimiento" estos valores coinciden, pero conforme λ aumenta van separándose y alejándose más y más. El aumento de valor de λ es también causa de que la pendiente de g en sus dos puntos fijos estables vaya creciendo progresivamente, hasta que al fin la función g –como ya le sucediera a su progenitora, la función f– llegará a su punto de ruptura (cuando su pendiente en los puntos x^*_1 y x^*_2 rebase los 45 grados) y cada uno de sus puntos de atracción se parta en otros dos, generando así su propio 2-ciclo local. (Debe quedar claro que los nuevos ciclos tan sólo son 2-ciclos con respecto a g; con respecto a f los nuevos puntos son elementos de un atractor de período cuatro. ¡No deben confundirse los papeles del f y g!) Estas dos escisiones se producirán exactamente en el mismo instante (es decir, para un mismo ajuste del control de λ), pues los valores de la pendiente de g en x^*_1 y x^*_2 son cada uno esclavo de otro. Esta especial posición del mando λ será llamada Λ_2; su valor es 0,86237...

Y en este punto, como en los chistes malos, vemos ya por donde van los tiros. Sin duda, el lector intuye que para una cierta posición Λ_3 los cuatro puntos del atractor de f se fisionarán simultáneamente, produciendo un atractor periódico constituido por ocho puntos, y que a partir de ahí esta conducta irá repitiéndose una y otra vez, duplicándose y reduplicándose los puntos del atractor cada vez que se alcanzan y rebasan ciertos valores umbrales de λ. Si tal ha sido su conjetura, está usted enteramente en lo cierto, y la razón subyacente es siempre la misma: las pendientes (idénticas) de todos los puntos fijos estables de cierta iterada de f alcanzan el ángulo crítico de 45 grados. En el caso de la primera fisión (para $\lambda = \Lambda_1$) se trataba de la pendiente de la propia f en el único punto x^*. La siguiente fisión se debió a que las pendientes de g en sus dos puntos fijos estables x^*_1 y x^*_2 alcanzaron ambas simultáneamente los 45 grados. Análogamente Λ_3 es aquel valor de λ para el cual la pendiente de $h(x) = f(f(f(f(x))))$ llega simultáneamente a 45 grados en los cuatro puntos fijos estables de h, y así sucesivamente. La ilustración superior muestra el escabroso aspecto de $h(x)$ para un valor de λ aproximadamente igual a 0,87.



Para $\lambda = 0{,}87$, la gráfica de **h(x)** *aún tiene más gibas que* **g(x)**

En la ilustración de la página siguiente vemos la localización sobre el eje x de los puntos fijos estables de f correspondientes a los valores Λ_1 a Λ_6 (para este último hay 32, algunos tan apiñados que es imposible distinguirlos a esta escala de dibujo). El trazado corresponde al preciso momento en que los valores fijos se hacen inestables, siendo cada uno a modo de una célula cuya división fuera inminente. Observemos la nítida pauta de la distribución de los puntos atractores. Basta un momento para apreciar que cada línea puede deducirse de la anterior por un esquema geométrico recursivo, en el cual cada punto se sustituye por dos "gemelos" situados bajo él. Cada condensación local de puntos se ajusta a un modelo que reproduce la pauta de la organización global precedente, aunque a escala más reducida (y en los cúmulos locales consecutivos, intercambiándose sus lados derecho e izquierdo). Por ejemplo, en la línea más baja hemos recuadrado en color un cúmulo de ocho puntos. Observemos que este grupo de puntos es como una miniatura de la configuración global dibujada dos líneas más arriba.

El descubrimiento de esta regularidad recursiva, que Mitchell J. Feigenbaum logró con el auxilio de una pequeña computadora, ha sido uno de los principales pasos dados en este campo. En particular, establece que para formar la $n + 1$-ésima línea a partir de n-ésima basta hacer que cada punto de la línea n-ésima dé a luz dos "gemelos". La nueva generación de puntos debiera quedar unas 2,5 veces más densamente empaquetada de lo que estaba la generación precedente. Enunciado con más precisión, la distancia entre los nuevos gemelos deberá ser α veces menor que la distancia entre su generador y el ge-

melo de éste, siendo α una constante igual a 2,5029078750958928485..., aproximadamente. Esta regla va haciéndose cada vez más exacta al ir creciendo el valor de n.

¿Qué puede decirse respecto de los valores Λ? ¿Tenderán asintóticamente al valor 1? Sorprendentemente, no es así. Los valores de Λ convergen raudos hacia un valor crítico especial, λ_c, aproximadamente igual a 0,892486418... Y su convergencia es de una regularidad notoria, en el sentido de que la distancia entre sucesivos valores de Λ se contrae en progresión geométrica. Con mayor precisión, la razón $(\Lambda_n - \Lambda_{n-1}) / (\Lambda_{n+1} - \Lambda_n)$ tiende hacia un límite finito, llamado δ por su descubridor, Feigenbaum, y "número de Feigenbaum" por los demás. Su valor aproximado es 4,669201660910299097...

En resumen, conforme λ va aproximándose a λ_c, para valores especiales de λ predichos por la constante δ de Feigenbaum, la población del atractor asociado a f se duplica y sus cada vez más numerosos elementos quedan situados sobre el eje x conforme a un plan recurrente sencillo, cuyo principal parámetro determinante es la otra constante de Feigenbaum, α.

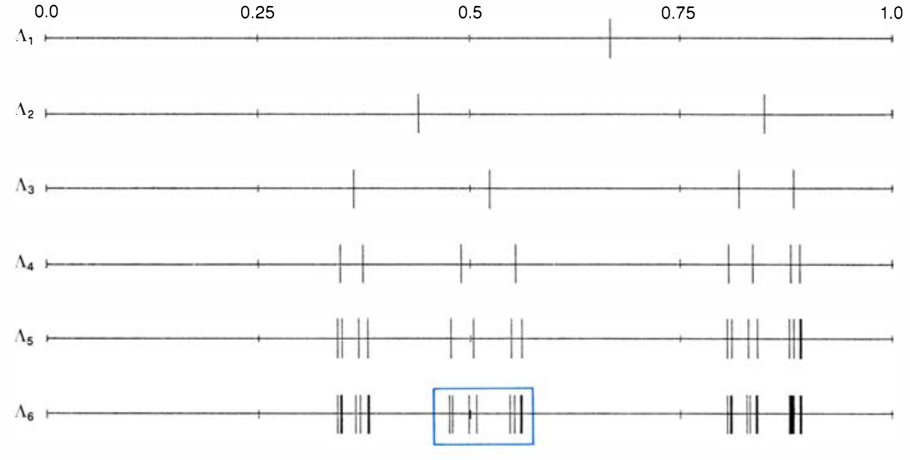
Y después, para λ mayor que λ_c –en el llamado régimen caótico– puede suceder que ciertos valores de las semillas generen, por iteración de f, órbitas no convergentes hacia sistemas atractores con número finito de puntos. Estas órbitas se llaman aperiódicas. Para la mayoría de las semillas, la órbita continúa siendo periódica; empero, su período resultará muy difícil de detectar, ante todo, porque será superlativamente grande, y en segundo lugar, porque la órbita será muchísimo más compleja que antes. Una órbita periódica de este tipo, en lugar de converger rápidamente hacia un atractor geométricamente sencillo, como antes, serpenteará por todo el intervalo [0,1] y su comportamiento será indistinguible del caos total. Se dice en estos casos que la órbita sigue un comportamiento ergódico. Además, semillas vecinas pueden muy bien, tras muy pequeño número de iteraciones, dar nacimiento a órbitas muy diferentes. En breve, cuando λ rebasa λ_c es mucho más razonable darle al fenómeno un enfoque estadístico.

Ahora bien, ¿qué relación puede existir entre funciones "combadas", la duplicación de períodos, el régimen caótico y nociones asociadas, con el estudio de las turbulencias en mecánica de fluidos, las fluctuaciones erráticas de las poblaciones en los sistemas "depredador-presa" o la inestabilidad de los modos láser? La idea central procede del contraste entre flujo laminar y flujo turbulento. En un fluido que discurre mansamente, el flujo es laminar (término de suave sonido que expresa que todas las moléculas del fluido se desplazan como automóviles por autopistas de varios carriles). Los rasgos esenciales de este fluir son (1), que cada vehículo sigue sensiblemente la misma trayectoria que su antecesor, y (2), que dos vehículos cercanos, tanto si se encuentran en el mismo carril como si corren por carriles paralelos, irán lentamente separándose uno de otro al transcurrir el tiempo, en proporción sustancialmente idéntica a la diferencia de sus velocidades, es decir, linealmente. También las moléculas de un fluido en régimen laminar comparten estos rasgos; en su caso, los carriles de flujo son las llamadas líneas o láminas de corriente.

En cambio, cuando un fluido se ve desgarrado por alguna fuerza exterior, tal apacible comportamiento se torna en turbulento, como vemos al romper las olas en la playa, o al echarle al café un chorro de leche. Hasta la propia voz "turbulento" tiene un sonido mucho más áspero que la dulce "laminar". No podemos ya seguir pensando en el tráfico por autopista. Las líneas de corriente se separan unas de otras y se enmadejan en las más complejas formas. En tales sistemas hay remolinos y revesas, y vórtices y torbellinos de muy distintos órdenes de magnitud, todos presentándose simultáneamente; por ello, dos puntos que un instante antes se encontraban muy cercanos pueden ahora hallarse en regiones totalmente distintas del fluido. La aparición de trayectorias en rápida divergencia caracteriza al flujo turbulento. La distancia entre dos puntos puede muy bien crecer exponencialmente –en vez de proporcionalmente– en función del tiempo; el coeficiente de que vaya afectado el tiempo en el exponente es el llamado número de Lyapunov. Al hablar de caos del flujo turbulento estamos refiriéndonos a esa rápida e impredictible separación de elementos vecinos. Y tal comportamiento es notablemente similar al de la rápida separación de dos órbitas (para valores de λ en el régimen caótico) cuyas semillas estuvieran inicialmente muy cercanas.

Lo cual sugiere que el escenario (así se le llama), que provoca que órbitas periódicas de bonito aspecto vayan progresivamente dando lugar a las embrolladas y caóticas órbitas de nuestra función parabólica, pudiera, por qué no, ser idéntico al escenario implícito en la transición al estado turbulento en fluidos y otros sistemas. Empero, aclarar cómo pueda establecerse la conexión entre ambos requiere explicitar con algún detalle su contexto. En particular, deberemos examinar brevemente la construcción de modelos matemáticos que describan la evolución espacio-temporal de un fluido o de algún otro sistema, como puedan ser la densidad de una población o la masa monetaria.

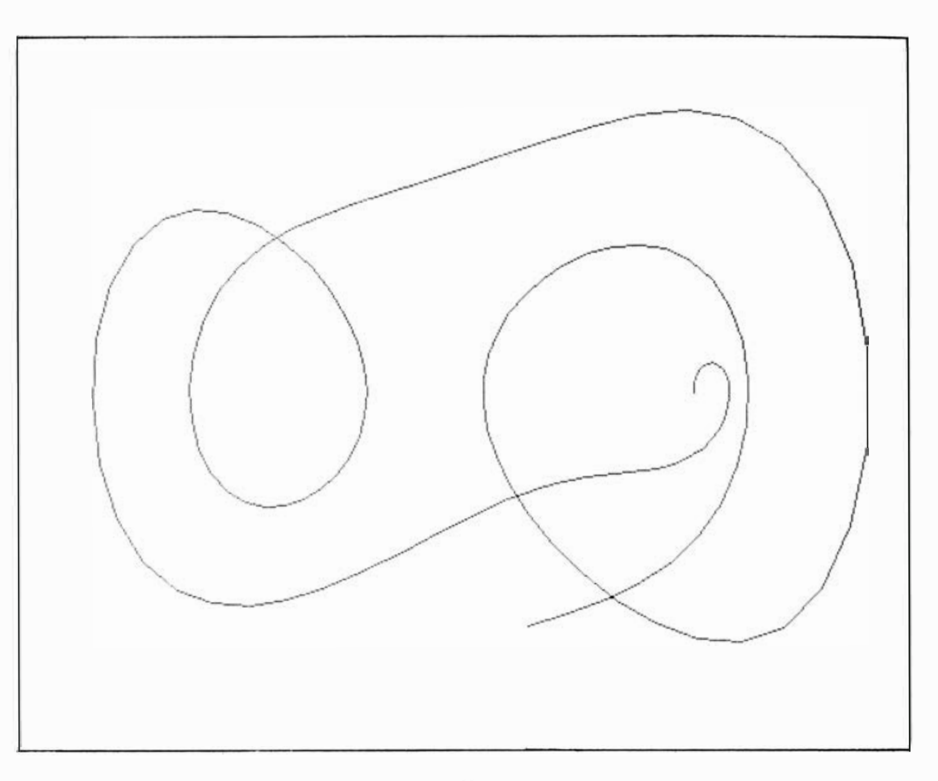
Hasta ahora, los modelos matemáticos que mejor describen este tipo de fenómenos se valen de ecuaciones diferenciales. En las ecuaciones diferenciales se establecen relaciones entre las tasas instantáneas de variación continua de ciertas magnitudes y los valores que en iguales instantes toman éstas y otras magnitudes que afectan al fenómeno. Además, cuando en ellas interviene el tiempo, se supone que discurre sin discontinuidades, regularmente, no como hacen los relojes eléctricos modernos, que saltan bruscamente de una cifra a otra. Una forma de visualizar las pautas



Los atractores estables se hacen inestables y "se fisionan" para una serie de valores del "mando Λ" (**de arriba abajo**)

definidas por ecuaciones diferenciales es imaginar un espacio multidimensional –que en ciertos problemas puede tener miles de dimensiones, o sólo dos o tres en otros– donde un punto va describiendo continuamente una curva. En cualquier instante, el movimiento del punto contiene toda la información relativa al estado del sistema físico; sus proyecciones sobre los ejes de coordenadas nos dan los valores de todas las magnitudes que determinan unívocamente el estado del sistema. Como es evidente, tal espacio –que se llama espacio de fases– habría de tener un enorme número de dimensiones para que uno solo de los puntos de la curva pudiera resumir la forma de una ola rompiendo en la playa. Por otra parte, en un sistema depredador-presa bastan sólo dos dimensiones: una variable, x por ejemplo, que exprese la población depredadora, y otra, y, que mida la población depredada.

Con el transcurrir del tiempo x e y van evolucionando, determinándose recíprocamente. Por ejemplo, si la población de depredadores es grande, disminuirá el número de posibles presas, mientras que una pequeña población de presas será incapaz de sustentar a un número grande de depredadores. En un sistema así, x e y definen un punto (x, y) que gira y valsea suavemente, describiendo una órbita continua sobre el plano. (El término "órbita" tiene ahora un sentido diferente al expuesto hasta aquí, donde las órbitas eran sucesiones de puntos individuales obtenidos por iteración de una función parabólica.) En la ilustración superior de esta página vemos una posible órbita del nuevo tipo; ha sido generada a partir de una ecuación diferencial conocida por ecuación de Duffing. Su aspecto recuerda el revoloteo de una mosca por la habitación, o mejor, la *sombra* que sobre una pared produciría la trayectoria de la mosca. En efecto, esta curva del plano *es* la sombra de una trayectoria tridimensional. La curva plana se corta varias veces a sí misma; la tridimensional, no. En el espacio de fases, la trayectoria del punto que describe el estado de un sistema no podrá jamás cortarse a sí misma. Ello es debido a que tal punto lleva codificadas en sus coordenadas toda la información sobre la evolución del sistema, incluida la futura, y por ello no pueden existir dos sendas que partan de uno y un mismo punto.



Una curva generada por la ecuación de Duffing que remeda la sombra de la trayectoria de una mosca

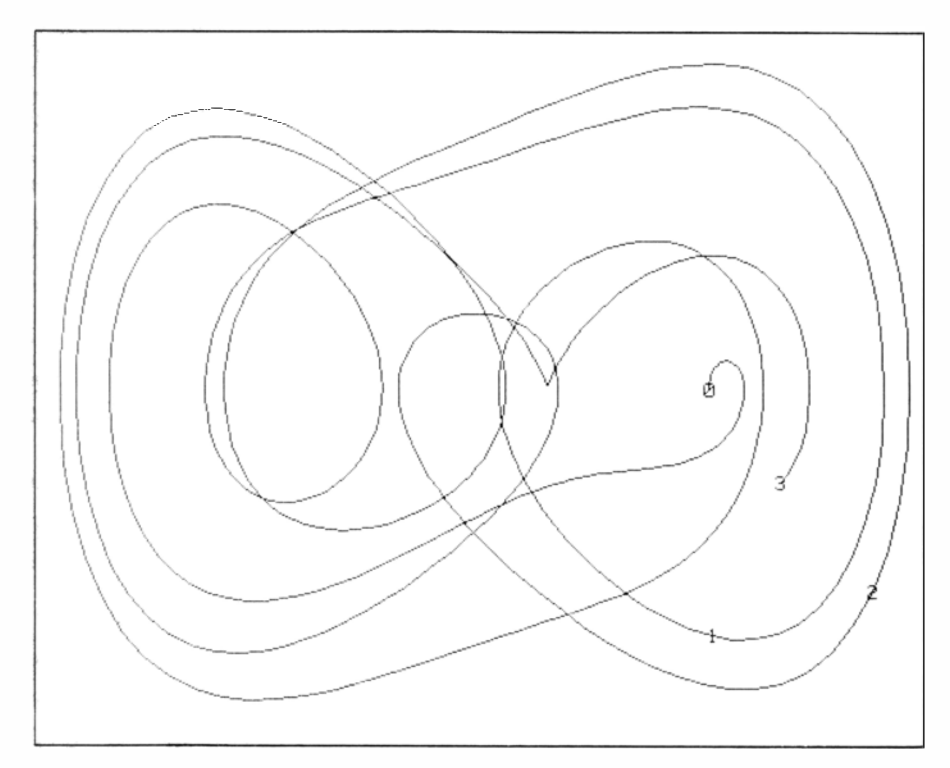
En particular, en la ecuación de Duffing hay una tercera variable, z, que hasta ahora no he mencionado. Si x e y son las poblaciones depredadora y presa, podríamos imaginar que z representa una influencia externa de variación periódica, como puede ser el azimut solar o el espesor de la nieve. Me tomaré ahora la licencia de hacer entrar en la escena de presas y depredadores a mi molesta mosca. Imaginemos una gran sala en la que la mosca vuela *periódicamente* de pared a pared. Supongamos que la mosca tarde un año en cruzar la habitación y retornar (se trata de una mosca tardona). Al volar la mosca de pared a pared, la sombra proyectada sobre una de ellas describe la curva que vemos en la ilustración superior. Si se diera el azar de que la mosca retornase a un punto de la sala por donde hubiera pasado en un viaje anterior, por el carácter periódico de su recorrido la mosca se vería condenada a realizar un bucle cerrado, siguiendo indefinidamente, una y otra vez, el camino que describió. Espero que con esto el lector se habrá hecho idea de cómo pueden ser las órbitas trazadas por un punto del espacio de fases que representa los posibles estados de un sistema dinámico regido por ecuaciones diferenciales.

Supongamos ahora que sea necesario establecer alguna clase de conexión entre estos sistemas y las órbitas *discretas*. ¿Cómo lograrlo? Bueno, no es preciso observar en cada instante los valores de x, y y z. Podríamos también ir tomando muestras periódicamente, con cierta frecuencia natural. En el caso de poblaciones animales, el año terrestre puede ser un valor muy natural del período. El azimut solar sigue la más estricta de las leyes periódicas, y la climatología de cada lugar se *esfuerza* en repetirse año tras año. Hay por tanto una sucesión natural de puntos aislados (x_1, y_1, z_1), (x_2, y_2, z_2)... destacables, a razón de uno por año. Es como si un estroboscopio –una lámpara de destellos– fuese relampagueando regularmente, "congelando" el vuelo de la mosca en ocasiones anuales especiales (¿por qué no en la última campanada de la noche de San Silvestre?). En todo otro momento, sus peregrinaciones por la sala permanecen invisibles. La ilustración de la página siguiente muestra una sucesión de puntos aislados, situados sobre la sombra de la trayectoria de la mosca, marcada con números que expresan cuándo fueron observados. Gradualmente, con el transcurrir de muchos "años", llegarán a acumularse suficientes puntos individuales como para empezar a formar una organización, de estructura reconocible por sí misma. Esta configuración de puntos es una "órbita" discreta, fuertemente relacionada con las órbitas discretas generadas por iteración en la parábola $f(x)$. En el caso parabólico teníamos una relación de iteración (o recurrencia) unidimensional muy sencilla: $x_{n+1} = f(x_n)$. La relación de recurrencia es bidimensional: $x_{n+1} = f(x_n, y_n)$ $y_{n+1} = f_2(x_n, y_n)$.

En este sistema de relaciones iterativas *acopladas*, los valores de salida de la n-ésima generación se reinsertan directamente como nuevos datos para f_1 y f_2, produciendo así la $n + 1$-ésima generación. Y esto, una y otra vez, generación tras generación. Evidentemente, en los problemas de mayor número de dimensiones el sistema tendrá un número de relaciones iterativas correspondientemente mayor; empero, el esquema en todos ellos es siempre el mismo: un punto multidimensional (x_n, y_n, z_n...) salta de una localización en el espacio de fases a otra cada vez que una variable discreta (n, que puede representar los progresivos saltos del tiempo, por intervalos discretos) es incrementada en una unidad.

Observemos cómo se ha escamoteado el escollo que suponía la variación continua del tiempo en las ecuaciones diferenciales. Lo hemos logrado concentrando la atención sobre la dependencia de los puntos respecto de sus predecesores de la "cosecha" anterior. Aunque, ¿habrá en todo sistema de ecuaciones diferenciales simultáneas un "período natural" que deba separar cada par de observaciones sucesivas? No siempre lo hay. Empero, en ciertas situaciones sí lo hay, y se da el caso de que así sucede siempre que se presentan fenómenos de turbulencia.

¿Y por qué razón? Todos los sistemas turbulentos son *disipativos*, lo que significa que disipan o degradan la energía de más "calidad", más utilizable, convirtiéndola en calor. En el caso del flujo hidrodinámico, ello se debe a la fricción interna del fluido, debida a su viscosidad, y en los otros sistemas que hemos estado examinando, por análogos abstractos del "rozamiento". La más conocida consecuencia de la fricción es que todo objeto en movimiento acabará por detenerse a menos que se le bombee energía para evitarlo. Ahora, si "dirigimos" un sistema disipativo excitándolo mediante una fuerza *periódica* (por ejemplo, al remover el café de una taza con un movimiento circular, periódico, de la cucharilla), entonces, evidentemente, el sistema no llegará a detenerse; tenderá más bien a adoptar alguna norma de "régimen". Tales estados estacionarios son órbitas estables, o con nuestra terminología, atractores del espacio de fases. Puesto que el sistema se ha excitado por una cucharilla en movimiento periódico, hemos definido al mismo tiempo una frecuencia natural con la que deberá destellar nuestros estroboscopio para "congelar" el estado del sistema, a saber, cada vez que la cucharilla pase en su revolución por cierta marca fija de la taza, el asa, por ejemplo. Así tendremos definido el "año" del sistema. De esta forma, la evolución continua puede sustituirse por una sucesión de instantáneas, siempre que estemos operando con un sistema disipativo excitado por una fuerza periódica. Y también así podremos sustituir las órbitas continuas por órbitas discretas, con lo que la iteración recupera su papel y retorna a escena.



Los números de la curva indican la posición de la mosca a intervalos regulares

Cuando la fuerza excitadora no tenga carácter periódico (por ejemplo, podría ser constante) hay todavía posibilidad de definir un período natural en tanto alguna variable del sistema oscile regularmente entre valores extremos. Bastaría que el "flash" destellase cuando tal variable alcanzara uno de sus valores extremos; la mosca seguiría así observándose en instantes discretos. Este tipo de representaciones discretas del movimiento de una partícula en espacios multidimensionales se llaman representaciones de Poincaré.

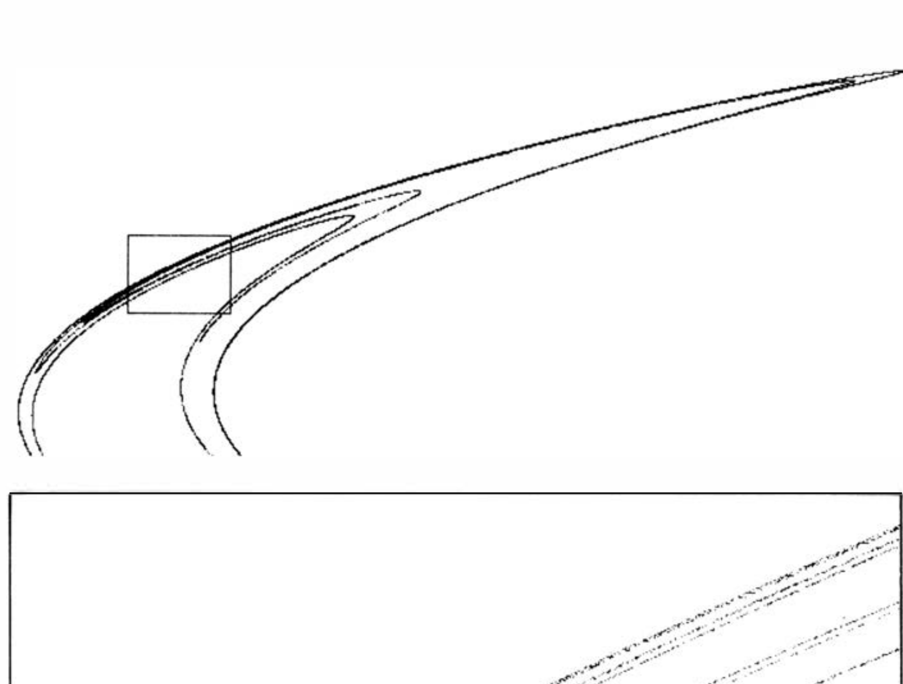
Todo este planteamiento carece de rigor, y malamente podrá convencer a ningún matemático, pero nos trae el aroma de la esencia del problema: estudiar sistemas de ecuaciones diferenciales mediante relaciones iterativas acopladas. Y este paso trascendente nos orienta hacia los recientes descubrimientos sobre la parábola.

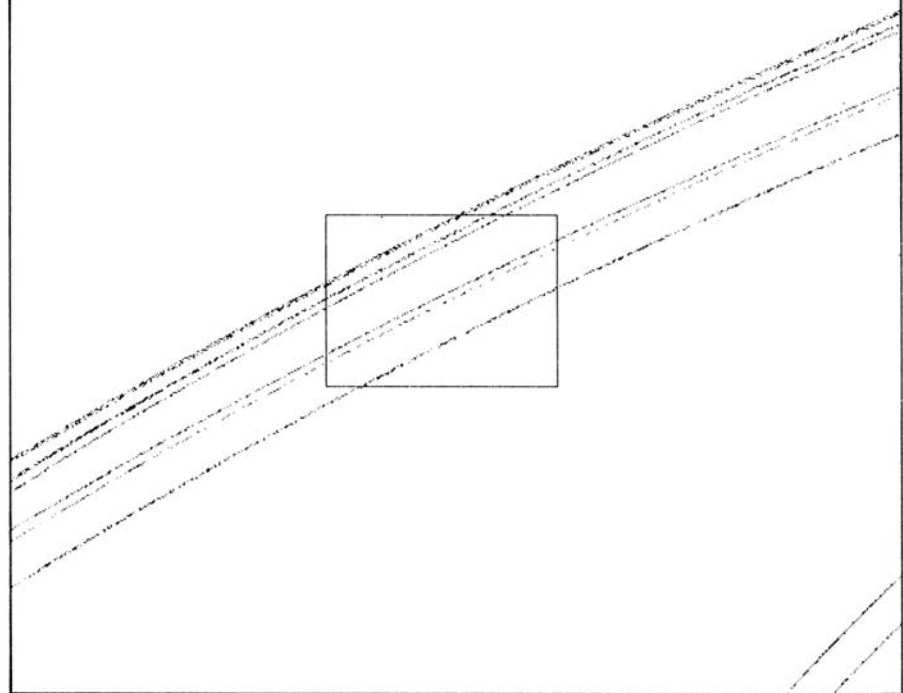
En 1975, Feigenbaum descubrió que sus números α y δ no dependían en realidad de la forma concreta que tuviera la curva definida por $f(x)$. Casi todas las curvas lisas convexas que alcancen su cima en el mismo punto que la parábola servirán perfectamente. Inspirado en la propiedad de universalidad estructural descubierta por Metropolis, Stein y Stein, Feigenbaum se sirvió de curvas sinusoidales, viendo atónito que reaparecían los mismos valores numéricos, hasta muchas cifras decimales, de los números que habían caracterizado la duplicación de períodos y el umbral del régimen caótico en la iteración de la parábola. Al igual que entonces, hay para las curvas sinusoidales un parámetro de altura, λ, y un conjunto de valores especiales de λ que convergen hacia un punto crítico, λ_c. Además, el umbral del régimen caótico, λ_c, está gobernado por los mismos números α y δ. Feigenbaum comenzó a sospechar que allí estaba ocurriendo algo de carácter universal. Dicho de otra forma, empezó a intuir que más importante que la propia f era el mero hecho de la iteración reiterada de f. Más aún, Feigenbaum conjeturó que tal vez f no tuviera influencia ninguna sobre el instante de aparición del caos.

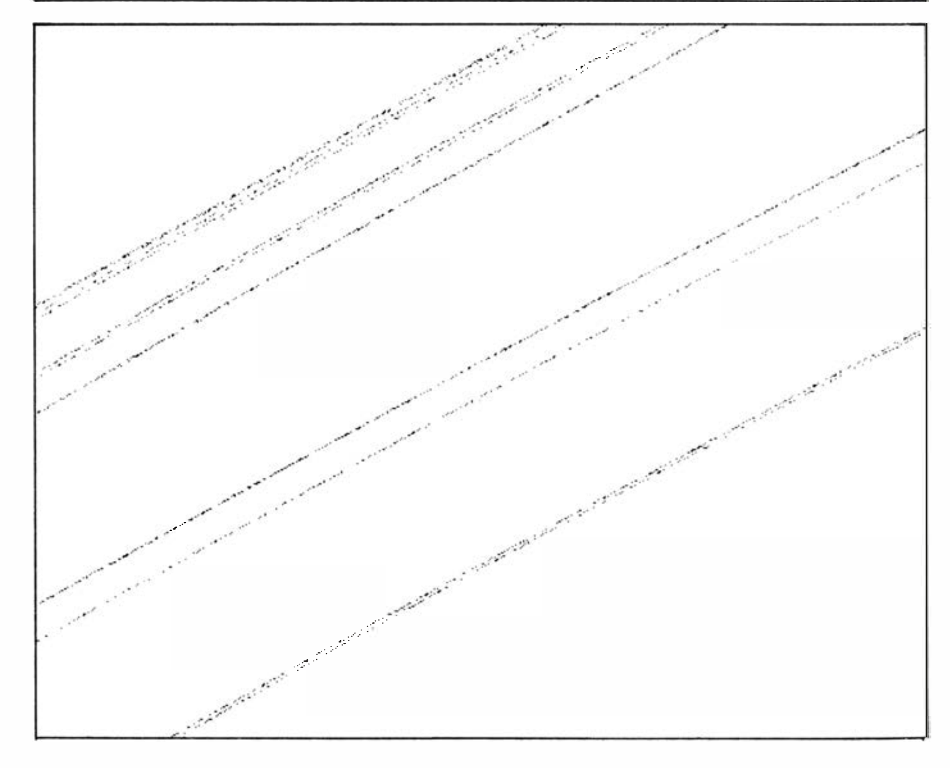
Evidentemente, las cosas no son tan sencillas. Feigenbaum descubrió pronto que lo realmente importante de f era la naturaleza de la cima situada en su centro. A largo plazo, el comportamiento de las órbitas depende tan sólo de un arco infinitesimal de la cresta de la gráfica y, en último extremo, únicamente de las propiedades de la curva en

el preciso punto donde alcanza su máximo. La forma que tenga el resto de la gráfica, incluso en la región cercana al pico, carece de importancia. Las parábolas tienen máximos de un tipo llamado "cuadrático"; también son cuadráticos los máximos de circunferencias, elipses y sinusoides. En realidad, la conducta que pueda mostrar en sus máximos locales una curva lisa tomada más o menos al azar será, de no darse coincidencias especiales, de tipo cuadrático. Por ello, el caso parabólico, más bien que curiosa excepción, parece ser quien da la regla. Este descubrimiento empírico de Feigenbaum, donde aparecen dos coeficientes escalares fundamentales α y δ, que caracterizan la formación del caos tras ir pasando por una sucesión de atractores, cada uno de período doble al de su predecesor, es manifestación de una propiedad universal de nuevo cuño, conocida como universalidad *métrica*, para distinguirla de la universalidad estructural anteriormente conocida. La validez del principio de univesalidad métrica ha sido demostrada (en el sentido matemático) para el caso unidimensional por Oscar E. Lanford III.

El descubrimiento de que las constantes de Feigenbaum no sólo se presentaban en atildados sistemas matemáticos, sino también en sistemas físicos donde se manifestaban embrollados fenómenos de turbulencia, fue todo un acontecimiento. Valter Franceschini, de la Universidad de Modena, Italia, adaptó la ecuación de Navier-Stokes (por la cual se rigen todos los flujos hidrodinámicos) para su simulación computarizada. A tal fin la transformó en un sistema de cinco ecuaciones diferenciales acopladas cuyas representaciones de Poincaré pudiera él estudiar después numéricamente con su ordenador. Descubrió primero que el sistema mostraba atractores cuyos períodos iban duplicándose conforme sus parámetros se acercaban a los valores donde se esperaba que surgiera el régimen turbulento. Desconocedor del trabajo de Feigenbaum, Franceschini mostró sus resultados a Jean-Pierre Eckmann, de la Universidad de Ginebra, quien inmediatamente le urgió para que retornase y determinase la razón de convergencia de los valores λ a los cuales se producía la duplicación de los períodos. Para asombro de ambos, aparecieron los valores α y δ de Feigenbaum con precisión de cuatro decimales, surgidos, diríase, de la nada. Por vez primera, un modelo matemático preciso de una turbulencia física reveló que su estructura estaba íntimamente relacionada con el







El atractor de Hénon, ampliado por dos veces, revela su infinita finura de detalle

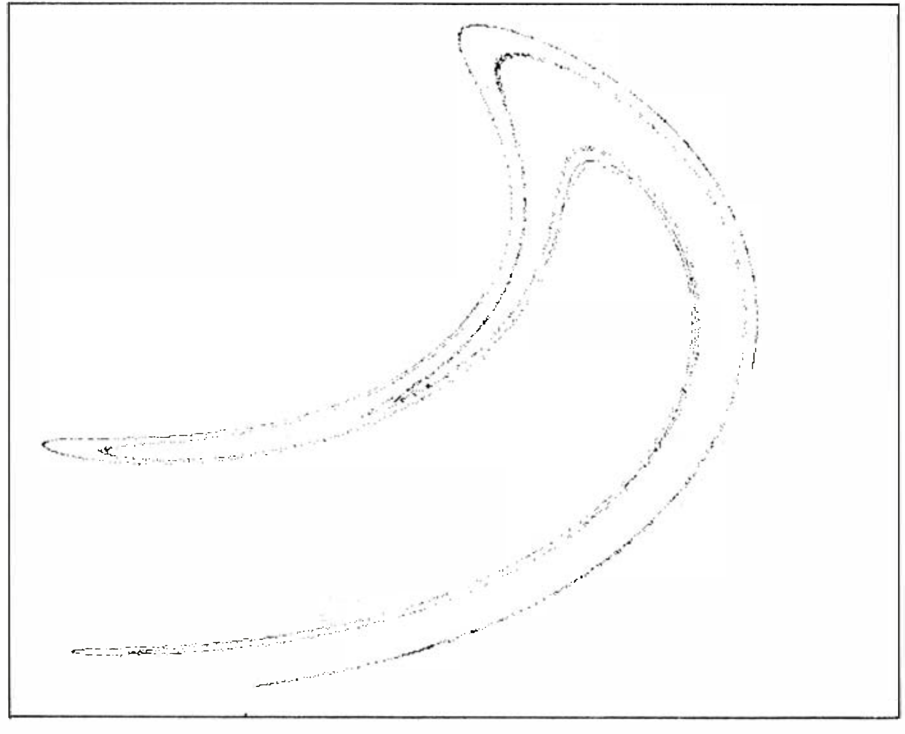
humilde caos que subyace oculto en la sencilla parábola $y = 4\ \lambda x\ (1 - x)$. En vista de estos resultados, Eckmann, Pierre Collet y H. Koch demostraron que, en la conducta de un sistema disipativo multidimensional sometido a excitación forzada, todas las dimensiones excepto una tienden a desaparecer transcurrido un tiempo suficientemente largo y, por ello, *es de esperar* que las características del comportamiento unidimensional –y concretamente, la universalidad métrica de Feigenbaum– hagan nuevamente aparición.

Desde entonces, los físicos experimentales han estado dejándose las pestañas en búsqueda, no de modelos computarizados, sino de sistemas físicos reales, que manifiesten la característica propiedad de duplicación de períodos. Se han observado conductas como las investigadas en ciertos tipos de corrientes de convección, pero hasta ahora las mediciones han sido demasiado imprecisas como para poder prestar apoyo suficiente a la idea de que la parábola pueda aportar en sí las claves para desvelar las turbulencias en sistemas físicos reales. Empero, resulta fascinante la idea de que, en realidad, lo único que importa es la iteración de un sistema disipativo de relaciones recurrentes acopladas, y que las propiedades "de detalle" de tales recurrencias puedan despreciarse plenamente si tan sólo nos interesa comprender la ruta que conduce al régimen turbulento.

Feigenbaum lo expone así: es frecuente ver en el cielo formaciones de nubes –un entretejido compuesto por una miríada de pequeñas bocanadas de vapor– extendiéndose de horizonte a horizonte. Es obvio que tal formación no se ha producido "por accidente"; alguna ley de la mecánica de fluidos ha tenido que determinar la formación. Mas, prosigue Feigenbaum, tendrá que ser una ley que opere a nivel superior, a mayor escala, que la ecuación de Navier-Stokes, pues ésta ha sido deducida a partir de volúmenes infinitesimales de fluido, y no "pedazos grandes". Nos parece que para comprender estas bellas formaciones nubosas tendremos que superar el análisis de "detalle" y servirnos de procedimientos que "no hilen tan fino", pero que destaquen más claramente los rasgos esenciales del flujo hidrodinámico. El descubrimiento de que la iteración genera universalidad –es decir, proporciona resultados independientes de las propiedades "de detalle" de la función o funciones que se están iterando– alumbra la esperanza de que bien pudiera estar en vías de aflorar un procedimiento de ese tipo.

Hasta ahora hemos hablado de atractores y turbulencias. Pero ¿qué son los atractores extraños? Disponemos ya de todas las nociones precisas para comprender esta nueva idea. Al construir, mediante un sistema de ecuaciones iterativas acopladas, un modelo matemático de un sistema disipativo (de dos o más dimensiones) sometido a excitación periódica, el conjunto de puntos captados por los sucesivos destellos de la lámpara estroboscópica va describiendo una forma o figura, que desempeña en este sistema el mismo papel que una de las órbitas sencillas de la parábola. Ahora bien, al operar en espacios de dimensión mayor que uno, se enriquece el abanico de posibilidades. Es posible tener un punto fijo estable: un atractor de período 1. En tal caso, al cabo del tiempo, para todos los destellos de la lámpara el punto que representa el estado del sistema se encontrará en el mismo lugar donde estuvo la última vez. También es posible tener un atractor *periódico*; de ser así, tras cierto número finito de destellos el punto habrá retornado a alguna de sus posiciones anteriores; la situación es análoga a la de los 2-ciclos, 4-ciclos, etcétera.

Pero cabe una nueva posibilidad: que el punto nunca retorne a una posición precedente del espacio de fases; los sucesivos relámpagos muestran que va saltando erráticamente en el seno de una región acotada del espacio de fases. Al cabo del tiempo, esta región podría empezar a tomar forma ante los ojos del observador, al proseguir el estroboscopio su periódico centelleo. En casi todos los casos hasta ahora observados, se ha podido apreciar un fenómeno sumamente inesperado: el punto errático va creando una delicada filigrana, que recuerda "los desvanecidos y fantásticos trazados de la helada en el cristal". Su delicada trama es de un tipo bastante especial, pariente cercana de las curvas "fractales" descritas por Benoit Mandelbrot en su libro *Fractals: Form, Chance and Dimension* [véase la sección "Juegos matemáticos" en INVESTIGACIÓN Y CIENCIA, febrero de 1977]. En particular, cualquier región del atractor, al ser ampliada, revela que su conformación es tan delicada como la de la imagen "ambiente" de la que se tomó. Es decir, hay una retrogradación infinita en el detalle, un interminable encaje de estructura en el seno de la estructura. Una de las primeras en descubrirse fue la llamada estructura de Hénon, que se reproduce en la página precendente. Está generada por la sucesión de puntos (x_n, y_n) definidos por las relaciones de recurrencia $x_{n+1} = y_n - ax_n^2 - 1$ e $y_{n+1} = bx_n$. En nuestro ejemplo, a es igual a 7/5 y b vale 3/10; los valores de las semillas son $x_0 = 0$ e $y_0 = 0$. El pequeño rectángulo de la ilustración superior se ha ampliado para mejor apreciar su detalle y, de ésta, se ha vuelto a ampliar otro



Un atractor extraño generado por la ecuación de Duffing

pequeño rectángulo en busca de más detalle. La configuración recuerda una autopista de tres carriles paralelos, cada uno de los cuales, al ser ampliado, se descompone en una nueva autopista, y así indefinidamente. Cada sección transversal perpendicular de esta autopista producirá uno de los llamados conjuntos de Cantor, que se construyen por un método recursivo ya clásico, que explicamos a continuación.

Se comienza por un intervalo cerrado [0, 1], por ejemplo. (El intervalo [0, 1] está formado por todos los números comprendidos entre 0 y 1. "Cerrado" significa que también sus extremos 0 y 1 pertenecen a él.) Ahora eliminamos un cierto subintervalo abierto central. Como el subintervalo abierto no incluye a sus puntos extremos, estos dos puntos permanecerán en el conjunto de Cantor que estamos construyendo. Por lo común, el subintervalo suprimido es el tercio central (1/3, 2/3), pero ello no es obligatorio. Hemos dejado dos subintervalos cerrados. Sometámoslos al mismo tratamiento, a saber, eliminar un subintervalo abierto central de cada uno de ellos. Repitamos el proceso ad infinitum. Al cabo quedaremos con una frágil estructura formada por una infinidad de puntos esparcidos a lo largo del segmento de partida [0, 1], como cuentas de rocío enfiladas en un hilo de araña. Empero, el conjunto de Cantor es infinito y no numerable; su densidad dependerá del procedimiento de supresión iterativa de intervalos que se haya utilizado. Tal es la complejidad de los conjuntos cantorianos; y cuando las secciones transversas de un atractor presentan distribuciones tan insólitas como éstas se dice que el atractor es extraño, ¡y con motivo!

Otro hermoso atractor extraño está generado por los puntos "estroboscópicos" 0, 1, 2, ... de la ilustración de la página 110. Como esta configuración procede de la ecuación de Duffing, se la conoce por atractor de Duffing; podemos verla a escala ligeramente ampliada en la ilustración de la página opuesta. Notemos su gran semejanza con el atractor de Hénon. Tal vez la universalidad vuelve a mostrar aquí su rostro.

Es interesante que para la parábola, en el valor crítico λ_c, el atractor de f se haga aperiódico, y pase a estar formado por infinidad de puntos. (Después de todo, es culminación de una sucesión infinita de duplicaciones del período.) Además, la disposición de esta infinidad no numerable de puntos del intervalo [0, 1] se ha construido mediante la regla recursiva de Feigenbaum, con ayuda de su constante α, como ya hemos visto en una ilustración precedente. No cuesta mucho aceptar que por esa razón este atractor concreto sea él mismo un conjunto de Cantor. ¡La fértil parábola nos ha proporcionado también un ejemplo de atractor extraño *unidimensional*!

En el régimen caótico correspondiente a casos más generales, de dimensión k mayor que 1, la predicción a largo plazo del camino que irá tomando un punto es completamente imposible. Dos puntos que ahora están casi en contacto se encontrarán en lugares sumamente alejados tras unos cuantos guiños de la luz estroboscópica; se dice por ello que el sistema presenta "gran sensibilidad a las condiciones iniciales", y esta sensibilidad es otro criterio definidor de los atractores extraños.

Hoy por hoy, nadie sabe por qué, cómo, o cuándo se podrán cosechar atractores extraños en los regímenes caóticos asociados a esquemas iterativos que representan sistemas físicos disipativos, pero sí parece cierto que ocupan lugar central en el misterioso fenómeno que llamamos turbulencia. David Ruelle, uno de los pioneros en promover este nuevo enfoque de la turbulencia, escribió: "Estos sistemas de curvas, estas nubes de puntos, evocan a veces galaxias y fuegos de artificio, y otras, floraciones, tan insólitas como intranquilizadoras. Hay todo un mundo de formas pendientes de análisis, de armonías aún por descubrir."

Un teórico de la biología, Robert M. May, en un artículo, hoy famoso, de 1976 donde pasaba revista a los hallazgos realizados hasta entonces en el campo, terminaba su trabajo con un ruego que deseo reproducir a continuación:

"Yo quiero... insistir en que a la gente debería siempre presentársele [la ecuación $y = 4\lambda x\ (1 - x)$] muy al comienzo de su formación matemática. Esta ecuación puede estudiarse fenomenológicamente por iteración, bien con una calculadora, bien a mano. En su estudio no se requieren elementos tan conceptualmente elevados como los del cálculo diferencial, siquiera sea elemental y, sin embargo, enriquecería mucho la intuición del alumno en lo relativo a sistemas no-lineales.

"No sólo en la investigación, sino también en el mundo ordinario –el mundo político y económico– todos nos defenderíamos mucho mejor si hubiera más personas que comprendieran que los sistemas no-lineales sencillos no tienen por qué poseer necesariamente propiedades dinámicas sencillas."

Taller y laboratorio

Delicias del estenoscopio y de su pariente, el antiestenoscopio

Jearl Walker

La deslumbrante diversidad y perfección de las cámaras fotográficas que hoy día se encuentran en el mercado ha ocultado la realidad de que pueden obtenerse fotografías muy aceptables sin otra cosa que un orificio diminuto, o estenope, interpuesto entre la película y el objeto de la fotografía: lo que constituye el fundamento de la cámara oscura, llamada también estenoscopio. Lo mismo cabe afirmar acerca del complemento óptico del estenope, o "antiestenope", que es una mácula, muy pequeña y circular, que se interpone entre la película y el objeto; así, puede hablarse del "antiestenoscopio". Para la fotografía estenoscópica voy a seguir las investigaciones de A. Connors, de la Universidad de Wisconsin, Madison, y de Matt Young, del National Bureau of Standards (Oficina Nacional de Pesas y Medidas). La novedosa y original idea de la fotografía antiestenoscópica procede de Adam Lloyd Cohen, de la Universidad de Loyola, Chicago.

La fotografía estenoscópica (estenopetografía) se basa en el paso de la luz a través de un orificio practicado en una pantalla opaca. Luego, la luz va a parar sobre un trozo de película donde reconstruye una imagen del objeto fotografiado. Referencias a las imágenes estenoscópicas se encuentran ya en Aristóteles; Leonardo da Vinci estudió sus principios y Lord Rayleigh las analizó formalmente. La sencillez constituye una de las muchas ventajas que ofrece la cámara oscura frente a las cámaras de objetivo refringente.

Al fotografiar un objeto con una cámara oscura, cada punto del objeto colocado frente a ella proyecta su propia manchita luminosa sobre la película (o papel fotográfico). La reunión de esas manchitas constituye la imagen que registra la película. Para que dicha imagen sea nítida, las manchitas contiguas no deben solaparse; habrán de ser, pues, lo más pequeñas posible.

En parte, el diseño de un estenoscopio descansa en la elección del tamaño del estenope y de una distancia entre dicho orificio y la película tales que las manchitas luminosas proyectadas sobre la película queden separadas y alcancen su máxima luminosidad. En principio, el tamaño del estenope puede ser ilimitado. Sin embargo, cuanto mayor sea su tamaño tanto más alejado de la película deberá colocarse el estenope y tanto mayor habrá de ser la película; de ahí que la práctica dicte un tamaño limitado para el orificio. Existe, asimismo, un límite teórico para la pequeñez del estenope.

Consideremos una cámara cuyo estenope se encuentre a una distancia razonable (varios centímetros) de la película. Supongamos que el estenope sea relativamente grande (demasiado con respecto a su distancia a la película) y que la cámara esté orientada hacia un manantial luminoso puntual y muy alejado. Los rayos luminosos procedentes de la fuente llegarán al estenope prácticamente paralelos entre sí y al eje geométrico que atraviesa el orificio por el centro y es perpendicular a la pantalla. Entonces, el radio de la manchita luminosa recogida en la pantalla será igual al radio del orificio y, por ser éste grande, la manchita en cuestión resultará grande también. En tal caso, cuando se fotografiaran un gran número de manantiales luminosos puntuales, las manchitas se solaparían en la película y no serían individualmente reconocibles cada uno de los manantiales.

Si se reduce el tamaño del estenope, se achica el tamaño de la manchita luminosa producida por cada fuente puntual. Ahora bien, la mejora de la situación por este procedimiento está limitada por el hecho de que el estenope acabaría siendo tan pequeño que la luz que lo atravesase se difractaría, para dar una figura de interferencia. De ese modo, una fuente puntual no crearía una pequeña mancha luminosa en la película, sino una figura circular compuesta de una mancha central luminosa, rodeada de anillos más tenues. Si menguáramos todavía más el tamaño del estenope, se agrandaría la figura de difracción, con la consiguiente pérdida de separación en la fotografía.

El radio óptimo del estenope es función de su distancia a la película; la relación puede mostrarse mediante un razonamiento en el que la luz se representase en forma ondulatoria. Imaginemos que se retiran el estenope y la pantalla; y consideremos una onda luminosa, procedente de un manantial puntual, que atraviesa el plano ocupado antes por la pantalla. Atendamos, en ese plano, a una familia de zonas circulares concéntricas con el eje geométrico de la cámara. Tales zonas pueden identificarse por su distancia a un punto situado en el centro de la película, punto que también está en el eje geométrico. La distancia entre la zona central y este último punto será la que mediaba entre el estenope y la película. La segunda zona está más alejada del punto central de la película en una semilongitud de onda, la tercera lo está en una semilongitud de onda más, y así sucesivamente.

Todas las zonas envían ondas luminosas al punto central; ahora bien, por tratarse de trayectos distintos, las ondas se interfieren a su llegada. Por ejemplo, la onda procedente de la segunda zona llega desfasada en media longitud de onda respecto a la onda procedente de la zona central. En este sentido, cuando las amplitudes de las dos ondas fuesen iguales, ambas se anularían entre sí. Desde luego, si las contribuciones tuvieran todas la misma amplitud, se anularían sin excepción en el punto central.

La verdad es que las amplitudes no son iguales, como se demuestra en exposiciones más rigurosas que ésta que puedo ofrecerles aquí, por lo cual la anulación es sólo parcial. La amplitud total de la onda luminosa en el punto central resulta ser la mitad de la que supondría la contribución de la zona central por sí sola. Como la intensidad luminosa depende del cuadrado de la amplitud, ello implica que la intensidad en el punto central es la cuarta parte de la que sería si sólo contribuyese la luz procedente de la zona central.

Pues bien, uno de los objetivos del estenope es obstruir todas las zonas, salvo la central. (Según afirman algunos investigadores las obstruye todas, excepto las dos primeras.) Un estenope de dimensiones óptimas permite que al punto central de la película sólo llegue luz de la zona central; con un estenope así, la manchita luminosa central será clara y pequeña, con la luz bien distribuida. Si el estenope no alcanza su tamaño óptimo, sólo contribuirá una parte de la zona central en la iluminación de la película, la mancha luminosa será más apagada y con la luz peor distribuida. Pero si lo rebasa en demasía, las zonas adicionales que abarca harán que

decrezca la claridad de la manchita, al tiempo que aumenta de tamaño.

Lo que se persigue, pues, no es que el estenope sea de un tamaño concreto, sino que la magnitud de éste y su distancia al punto central de la película guarden determinada relación. Así, cuando el objeto a fotografiar se encuentra relativamente alejado, el radio óptimo del estenope viene a coincidir, aproximadamente, con la raíz cuadrada del producto de la longitud de onda de la luz por la distancia entre el orificio y la película.

A partir de la relación mencionada cabe definir una distancia focal de estenope. Este actúa como un objetivo, en el sentido de que proporciona una imagen concentrada de un objeto. Su distancia focal es aproximadamente la longitud de onda de la luz partida por el cuadrado del radio del orificio. Cuando la película diste del estenope una longitud igual a la distancia focal, la manchita luminosa en la película será reducida y clara, pues sólo entonces la zona central llena por completo el estenope y contribuye con su luz al punto central.

Supongamos que el objeto esté cercano. Si hubiéramos de fotografiarlo con un objetivo refringente, podríamos calcular la distancia correcta entre éste y la película aplicando la fórmula llamada de las lentes delgadas, que establece que la inversa de la distancia entre objetivo y película debe ser igual a la inversa de la distancia focal del objetivo menos la inversa de la distancia al objeto. Esta relación mantiene su validez para los estenopes, con tal que la distancia focal se defina de la forma indicada antes; ello permite enfocar una cámara oscura y obtener así una fotografía de la mejor nitidez.

Si el objeto está alejado, la mejor posición de la película es la correspondiente a la distancia focal. Si nos desplazamos hacia el objeto, acortando la separación entre el mismo y el estenope, deberemos incrementar la distancia entre la película y el estenope, a fin de mantener la nitidez óptima en la imagen. Una corrección tal quizá no resulte muy viable, ya que en los estenoscopios suele ser fija la distancia entre estenope y película. En cuyo caso vale la pena sustituir el estenope por otro menor, de suerte que disminuya la distancia focal.

En la práctica no se efectúa ninguna de esas correcciones, dado que la nitidez de la fotografía suele ser aceptable, aun cuando el tamaño del estenope y la distancia entre éste y la película no cumplan la relación óptima. Y si se fotografía una escena cuyos objetos se encuentran a distancias de la cámara que varíen entre límites amplios, en su mayoría aparecerán en la fotografía aceptablemente enfocados. Esta gran profundidad de campo es característica del estenoscopio.

A partir de lo expuesto podemos calcular el tamaño conveniente del estenope, o bien la distancia entre éste y la película, una vez escogido uno de ellos. Pero, ¿cuál de esos datos se elige para calcular el otro? La respuesta la dicta el sentido práctico. En efecto, nadie desea un estenoscopio de varios metros de longitud. Además, queremos una foto perfecta, en la que se aprecien los detalles que nos ofrecen la contemplación directa de la escena. Aquí, precisamente en la nitidez de la imagen, se



Estenopetografía de Kenneth A. Connors

encuentra el punto de partida para empezar a establecer las condiciones de la cámara.

El límite hasta el que el ojo humano es capaz de percibir con nitidez se llama poder separador, magnitud angular que se expresa en radianes. Supongamos que nuestro campo visual abarca dos puntos. En tanto que el ángulo que subtiendan sea superior a determinado valor mínimo, 0,001 radianes aproximadamente, podremos discriminarlos. Cuando dicho ángulo sea menor, sólo veremos un objeto único y borroso. Por ejemplo, dos puntos contiguos, separados un milímetro entre sí y a un metro de nosotros, se encuentran precisamente en el límite de nuestra capacidad para verlos con nitidez. En tal caso bastaría una cámara con ese poder separador; aumentarlo no valdría de nada.

A título demostrativo, supongamos que el fotograma definitivo haya de tener las mismas dimensiones que la película y que vamos a contemplarla a una distancia igual a la existente entre el estenope y la película. Propongámonos fotografiar dos fuentes luminosas puntuales y continuas, cuya separación angular corresponde al poder separador del ojo humano (0,001 radianes). La cámara deberá proyectar sobre la película dos manchitas que apenas se toquen, o que se solapen muy poco. Entonces, al observar la fotografía, podremos verlos como manchitas apenas separadas. El ángulo que formen esas manchitas en nuestro campo visual puede calcularse dividiendo el diámetro del estenope por la longitud de onda de la luz. Sea ésta de 500 nanómetros (aproximadamente el centro del espectro visible); si el ángulo que mide nuestro poder separador es 0,001 radianes, el radio del estenope habrá de cifrarse en 0,25 milímetros.

Determinado este valor a través de la fórmula antes referida se deduce que la



Fotograma gran angular obtenido por John M. Franke con una cámara oscura dotada de una semiesfera de vidrio detrás del estenope

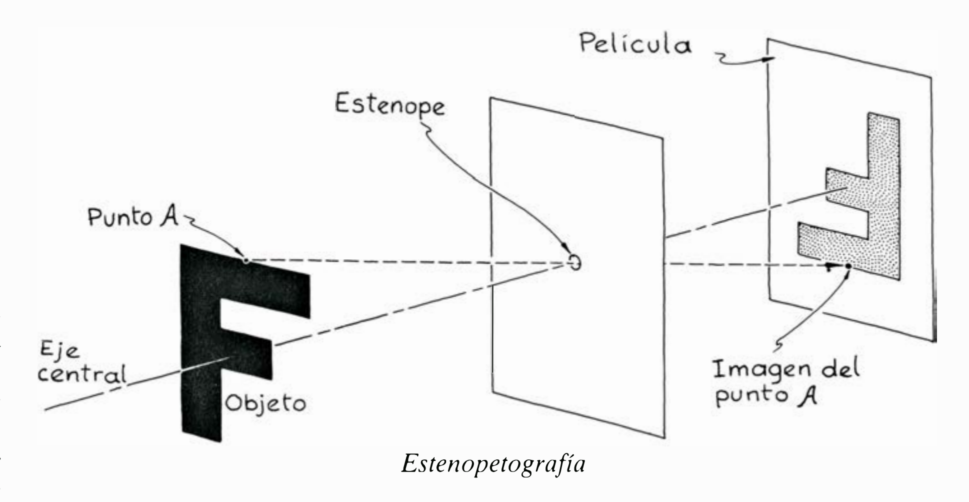
distancia óptima entre el estenope y la película es de 12,5 centímetros. Si duplicáramos el radio del estenope y, de acuerdo con ello, corrigiéramos la distancia de la película al orificio, la nitidez de una foto hecha con esa cámara tendría una calidad doble. Pero si no cambiamos también el tamaño del fotograma y la distancia a la cual lo contemplamos, no habrá manera de advertir la mejora. Por otro lado, la cámara habrá de tener aquí 50 centímetros de longitud (para que la distancia entre el estenope y la película sea la correcta) y se necesitará una película mayor para recoger toda la luz procedente del estenope. Es evidente que la cuantía de la mejora no vale la pena.

Cuando el estenope sea mayor de lo debido, la nitidez perderá calidad, con el inconveniente adicional de que aparecerán detalles falsos en el fotograma. Este efecto, que se llama separación espuria, resulta del solapamiento de las imágenes de varios objetos contiguos. En la ilustración superior de la página siguiente aparece la explicación de Young acerca de la separación espuria de tres barras verticales.

La mayoría de los sistemas de lentes producen una distorsión lineal en las imágenes registradas sobre película. Por ejemplo, un objeto cuadrado puede aparecer como si sus lados fuesen levemente curvos. La mayoría de las cámaras modernas incorporan dispositivos para corregir este fallo. Pero una de las ventajas del estenoscopio es que se encuentra prácticamente libre de distorsión lineal.

El estenoscopio presenta, empero, varios tipos de aberración, incluida la cromática. En efecto, como el radio óptimo del estenope (y, por ello, su distancia focal) depende de la longitud de onda de la luz, la cámara sólo puede optimizarse para una longitud de onda. Entonces, podrá conseguirse la mejor nitidez para esa longitud de onda, pero la que se consiga para las demás longitudes de onda que integran la luz blanca será de peor calidad.

Con película en color, los bordes de las imágenes aparecen borrosos; a veces se aprecian incluso coloreados. Con película en blanco y negro sólo se observa un emborronamiento de los bordes. Para eliminar la aberración cromática suele emplearse película en blanco y negro con un filtro de color colocado ante el estenope. El tamaño del estenope y la distancia de la película al mismo se optimizan para la longitud de onda que deja pasar el filtro. Se eliminan los demás colores y los bordes de la imagen aparecen menos borrosos por aberración cromática.



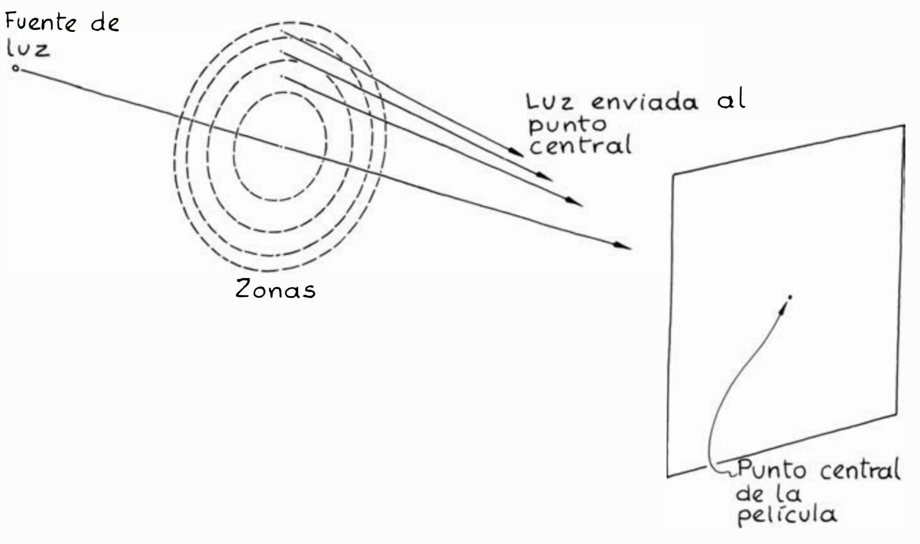
Estenopetografía

Otra aberración que presentan las cámaras oscuras es el astigmatismo, que se manifiesta cuando el objeto fotografiado se encuentra fuera del eje central del estenope. Aquí ocurre que el estenope visto desde el objeto tiene forma elíptica, no circular, y si el objeto es una fuente luminosa puntual, sobre la película se proyectará una manchita elíptica. Además, el lugar de la película sobre el que se proyecta la manchita no estará a la distancia correcta del estenope. Si el centro de la película se coloca a la distancia correcta del orificio, todos los demás puntos de la película quedarán demasiado lejos del orificio, lo que significa que únicamente se obtendrá una nitidez óptima en el centro.

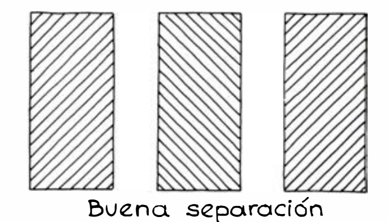
Un problema más grave de las cámaras oscuras es su reducida aptitud para captar luz. Debido a su apertura, habitualmente muy pequeña, se necesitan exposiciones prolongadas. Por ejemplo, si la película está a la distancia focal del estenope y esa distancia es de unos pocos centímetros, el número *f* de la cámara se cifrará aproximadamente en 200. Aunque la pequeña apertura hace que el sistema sea lento, tiene a su favor la gran profundidad de campo de la cámara.

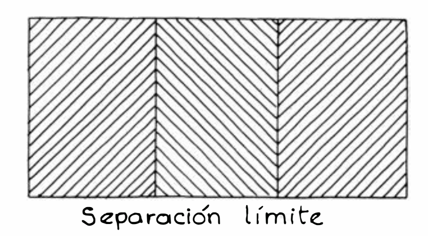
Por varias razones, la intensidad luminosa proyectada sobre la película no es nunca uniforme. Supongamos que van a fotografiarse dos manantiales luminosos puntuales, uno en el eje central y otro fuera de él. La luz procedente de la fuente puntual situada fuera del eje se encuentra con un estenope aparentemente elíptico; en consecuencia, por el orificio atravesará menos luz procedente de ese punto que procedente del situado en el eje. Asimismo, la luz que genera la manchita de fuera del eje ha de recorrer la distancia mayor para llegar a la película y, por tanto, se extiende más, alcanzando así la película con menor intensidad. Por otro lado, esa luz llega a la película bajo un ángulo que esparce aún más la exposición sobre la superficie de la película, reduciéndose más la intensidad.

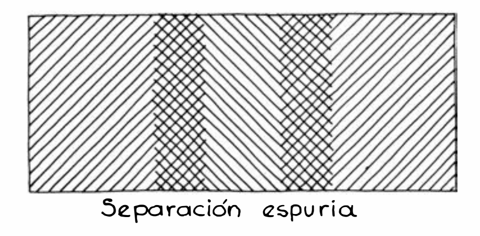
La limitación del campo responde a otra nueva razón: no todo objeto suficientemente separado del eje central refleja la luz hacia la película, salvo que ésta sea muy ancha o se coloque bastante cerca del estenope. Se ha intentado siempre salvar esa dificultad arri-



Zonas cuya luz contribuye al centro de la película







Tipos de separación

mando la película hacia el estenope, para que pueda obtenerse una fotografía de amplia angulación. Pero este truco tiene un fallo: reduce la nitidez de la foto, ya que la película no estará entonces a la distancia del estenope correspondiente a la nitidez óptima.

Otro procedimiento para aumentar el campo es preparar un soporte para la película en forma de semiesfera, centrada en el estenope. De ese modo, toda la luz que penetre por éste alcanzará la película, incluso la procedente de objetos situados casi a 90 grados respecto al eje central. Otra consecuencia sería un menor debilitamiento de la exposición, ya que la luz incidiría siempre perpendicularmente sobre la película, y mejoraría la nitidez para los objetos separados del eje central, al estar todas las porciones de película a la misma distancia del estenope. Desgraciadamente, la construcción de un soporte semiesférico no resulta muy viable. Una solución de compromiso suficiente podría ser un soporte cilíndrico.

John M. Franke, del Centro de Investigaciones Langley, de la NASA, resolvió esa dificultad de un modo ingenioso. Puso una semiesfera de vidrio inmediatamente detrás del estenope de la cámara, en la cual el soporte de la película era una placa plana normal. Cuando la luz atravesaba el estenope y penetraba en el vidrio, se refractaba. Entonces, la totalidad del campo, que abarcaba 180 grados, se reducía a un cono de luz de 84 grados de abertura. Luego, cuando la luz emergía del vidrio, caía perpendicular a su superficie y, por ello, no se alteraba el ángulo del cono de luz. Esta reducción angular de 180 a 84 grados le permitió a Franke ubicar la película a una distancia del estenope conveniente y, aun así, obtener un fotograma gran angular con un campo de casi 180 grados.

Franke empleó una semiesfera de 25,4 milímetros de diámetro, construida con vidrio BK-7, cuyo índice de refracción es 1,5 aproximadamente. Aunque el diámetro no incide de una manera crítica, con vidrios de diferentes índices de refracción se consiguen diferentes resultados. Puede resultar entretenido ensayar con otros vidrios, e incluso con plásticos de buena calidad. Si se desea un campo de 180 grados, se encontrarán imágenes algo distorsionadas hacia los bordes de la foto.

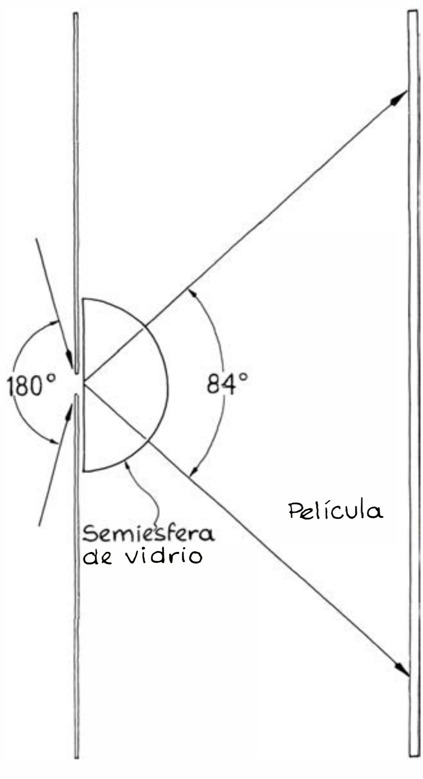
Hay varias formas de construir un estenope. Préstese atención al orificio, circular y con el borde liso. Young ha conseguido estenopes perfectamente limpios sirviéndose de láminas de latón de 50 micrometros de espesor (como las que se emplean para hacer suplementos espaciadores, etc.). A tal fin, monta una aguja de coser en una fresadora y luego, con el avance vertical de la máquina, impulsa la aguja a través de la delgada lámina de latón, bajo la cual coloca un bloque de plomo recién pulido para evitar que se deforme. Tras eliminar las rebabas del borde del orificio, lo escaría con la punta seca de un compás y vuelve a limpiarlo.

Las láminas de latón que emplea Connors miden de 25 a 50 micrometros de espesor. No se recomiendan plaquitas más gruesas, si se quiere evitar que el orificio se convierta en un cilindro demasiado largo en cuya superficie interior se produzcan más reflexiones de los rayos luminosos. Sobre cartón consistente o madera blanda pulida coloca un trozo de lámina cuadrado. Luego, valiéndose de la punta seca de un compás, Connors practica una pequeña embutición en el trozo de lámina, con cuidado para que la punta no la traspase por completo. Seguidamente, da la vuelta a la pieza y fricciona con tela de esmeril fina el conito formado en el reverso de la embutición. Estas operaciones las repite, 15 veces si es preciso, hasta que aparece un orificio suficientemente ancho para que por él pase la caña de la punta del compás, cuyo diámetro ha medido previamente con un microscopio de retículo graduado, de modo que ya sabe cuál es el tamaño del estenope. Cuando quiere un estenope de diámetro menor que la punta de compás de que dispone, detiene el proceso de agrandamiento antes de que toda la punta penetre en el orificio.

Conseguido el estenope, Connors encola la lámina a un soporte hecho de placa de latón más gruesa (de unos 125 micrometros de espesor), de modo que el estenope quede centrado en un orificio de unos seis o siete milímetros taladrado en la pieza más gruesa. El lado de este conjunto que ha de encararse a la película lo pinta de negro mate para disminuir las reflexiones luminosas en el interior de la cámara. Hay quien opina que también debe ennegrecerse la superficie interna del estenope, pero Connors se opone a estropear la simetría del orificio construido; sólo pinta hasta una profundidad de uno o dos milímetros.

Connors advierte que los estenopes deben mantenerse libres de polvo y, así, guarda cada uno en un saquito de plástico hasta que lo necesita. Periódicamente los examina al microscopio para comprobar si el polvo ha dañado su simetría.

Estos objetivos estenoscópicos pueden montarse, prácticamente, en cualquier tipo de caja hermética a la luz. Yo mismo he visto estenoscopios hechos con cajas de cereales. Trabajando en un cuarto oscuro, el fotógrafo monta un trozo de papel fotográfico en la parte posterior de la caja y desliza la tapa. Luego, para evitar que penetre luz prematuramente en el interior de la caja, coloca un trozo de adhesivo negro so-



Montaje de Franke para fotos gran angulares

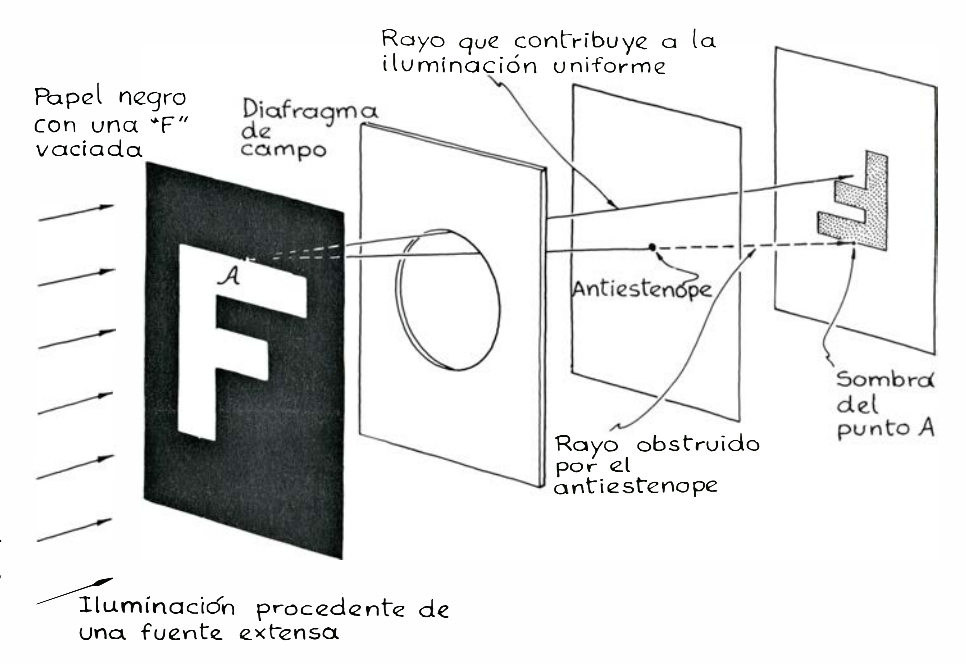
bre el estenope. Cuando todo está listo, efectúa la exposición arrancando el adhesivo del orificio y volviendo a colocarlo después. Una cámara así funciona, desde luego, como estenoscopio, pero tiene dos inconvenientes: sólo puede tomarse una foto cada vez que se saca del cuarto oscuro, y puede emborronarse la foto con la sacudida que resulta de la retirada y reposición del adhesivo.

Por mi parte, prefiero el procedimiento que apunta Young. Sobre la base de su cámara de 35 milímetros monta un tubo telescópico, que existe en el mercado para la mayoría de las cámaras de objetivo desmontable; en su extremo externo sujeta el objetivo estenoscópico. A falta de tubo telescópico, aproveché un tubo postal de cartón que fijé a la base de mi cámara con varias capas de cinta adhesiva negra. La ventaja de este tipo de estenoscopio es que permite la impresión de un carrete de película entero. Da la coincidencia de que mi cámara es un modelo reflex de objetivo simple, por lo que pude observar una imagen tenue de la escena antes de tomar cada foto.

En la estenopetografía, la luz atraviesa un orificio para generar una imagen: en la antiestenopetografía de Cohen, un antiestenope, u obstáculo, proyecta una imagen negativa del objeto. Este dispositivo constituye el complemento óptico del estenope. Aquí, la pantalla y el orificio están sustituidos por un pequeño obstáculo de sección circular. Así, la luz que hubiera pasado por el estenope queda obstruida, y la luz que hubiera sido obstruida por la pantalla llega a la película, dando una imagen negativa. El fotograma antiestenoscópico final es similar al estenoscópico, con una salvedad: las zonas claras y oscuras aparecen intercambiadas.

En las imágenes que proyectan los antiestenopes de Cohen no interviene la difracción de la luz, pues los antiestenopes son demasiado grandes para originar figuras de difracción importantes. La imagen la crea sencillamente la obstrucción de los rayos luminosos procedentes del objeto. Toda manchita que aparece en la película registra la sombra de una porción del objeto que se encuentra en la intersección del objeto con la recta definida por la manchita y el antiestenope.

Los fotogramas obtenidos con antiestenopes tienen peor contraste que los fotogramas estenoscópicos, porque los dispositivos antiestenoscópicos permiten que a la película llegue casi toda la luz procedente de la escena. De esa luz, la mayor parte es una iluminación uniforme que no sirve para nada y que



Montaje antiestenoscópico de Adam Lloyd Cohen

simplemente reduce el contraste de la imagen. Es el resto de la luz, la fracción no uniforme, la que transporta la información acerca del objeto. El contraste mejorará si hubiera alguna forma de disminuir la iluminación uniforme, o aumentar la proporción de luz no uniforme responsable de la información acerca del objeto.

La desigualdad entre ambas iluminaciones puede aumentarse si el antiestenope se acerca a la película; ahora bien, lo mismo que ocurría en la estenopetografía, esa medida rebaja la nitidez del fotograma. No obstante, Cohen afirma que con gusto sacrifica algo de nitidez para conseguir un contraste suficiente que deje reconocer la imagen en la foto.

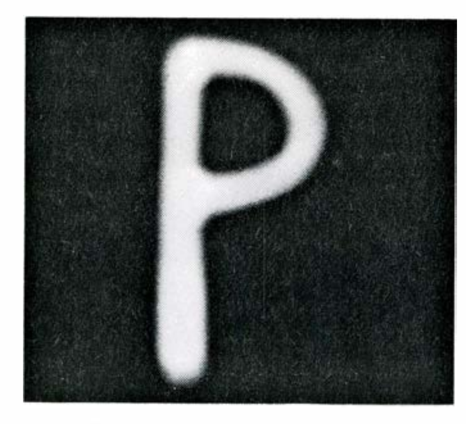
Parte de la iluminación uniforme proviene de porciones de la escena no importantes para el fotograma. Para reducir esta iluminación irrelevante, Cohen sitúa un diafragma de campo (una pantalla con un orificio mayor que el antiestenope) delante del antiestenope. El orificio es lo suficientemente ancho para que las porciones extremas del objeto iluminen los bordes de la película, y lo suficientemente pequeño para impedir que el resto de la escena llegue a la película.

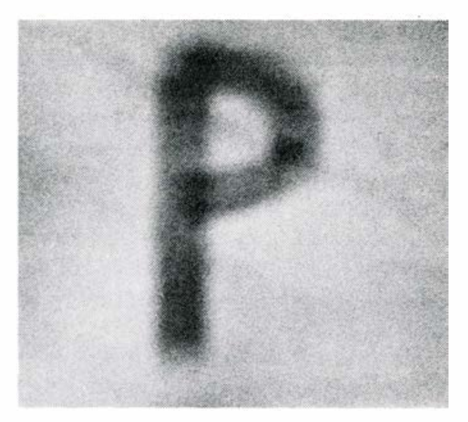
El coloreado de las imágenes antiestenoscópicas puede ser sorprendente. Cuando se fotografía un breve conjunto de objetos de diferentes colores, la imagen de cada uno suele aparecer con un color diferente al suyo. Este cambio depende del color que da la combinación de los colores de los objetos del conjunto. Si la combinación da blanca, cada color del conjunto cambia a su complementario en la fotografía; un objeto rojo dará una imagen cuyo color será la sustracción de blanco menos rojo (puesto que el antiestenope bloquea el rojo procedente del objeto). Por consiguiente, el color de la sombra es el azul verdoso llamado cian, complementario del rojo. Correspondientemente, un objeto verde crea una sombra magenta.

Algunas de las propiedades de las cámaras oscuras las comparten por igual los antiestenoscopios. Su campo es grande, el ajuste del aumento se consigue variando la distancia entre el estenope y la película y no existe distorsión lineal. En los antiestenoscopios puede evitarse el astigmatismo si el antiestenope se construye esférico, pues así toda la luz, procedente del objeto en dirección a la película, se ve interceptada por un obstáculo de sección circular, aun cuando el objeto esté muy separado del eje central de la cámara.

Otra diferencia entre ambos tipos de fotografía: una serie de estenopes alineados entre el objeto y la película no producen fotogramas, lo que sí ocurre en una serie de antiestenopes. Ello se debe a que las pantallas donde se perforan los estenopes impiden que la luz se proyecte sobre la sucesión de estenopes más cercanos a la película; por el contrario, los antiestenopes se interfieren muy poco. La ilustración inferior de la página siguiente es una fotografía obtenida por Cohen con una serie de antiestenopes que colocó entre un anillo brillante y la película. Cada antiestenope produce su propia imagen en negativo del anillo.

A quienes pueda interesar la anties-

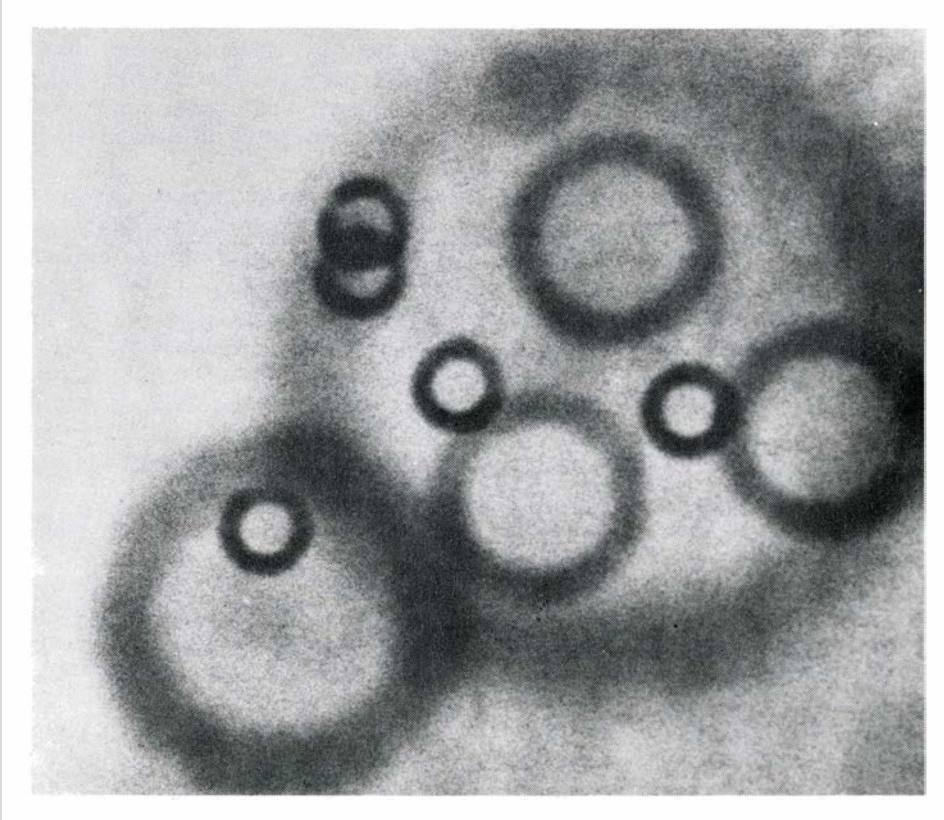




*Estenopetografía (***izquierda***) y antiestenopetografía de una P recortada en papel obtenidas por Cohen*

tenopetografía, Cohen les ofrece las sugerencias siguientes. Como antiestenope se utilizará una manchita redonda de pintura negra depositada sobre un trozo de vidrio o acetato limpio; no es crítico el tamaño de la manchita y, en lugar de ella, puede pegarse un redondel, como los que usan para rotular y expenden las tiendas de artículos para oficina. Cohen recomienda que el redondel no sea muy pequeño, que permita conservar el contraste del fotograma. Se aconseja una escena con mucho contraste para que el fotograma también lo muestre. La experimentación en antiestenopetografía puede comenzarse recortando figuras en papel negro opaco e iluminándolas por detrás con una fuente de luz extensa.

Sobre las fotografías estenoscópica y antiestenoscópica hay muchas más cosas que saber además de lo que yo les he explicado aquí. En *Optica Acta* no tardará en aparecer una descripción del trabajo de Connors bajo el título "Antipinhole Imaging" (Imágenes con antiestenoscopios). Parte de los trabajos más exhaustivos acerca de la estenopetografía, tanto experimentales como teóricos, pueden encontrarse en una serie de artículos publicados por Connors en *Interest*, una revista editada por él y que se le puede solicitar a School of Pharmacy, University of Wisconsin, Madison, Wis. 53706. Sus últimos artículos han versado acerca de las condiciones de nitidez y definición óptimas, el calibrado de una cámara para regular el contraste y la relación entre estenoscopios y óptica de las redes zonales.



Anillo brillante, fotografiado por Cohen a través de una serie de antiestenopes

Libros

Sobre metodología y fundamentos en matemáticas, física y ecología

Javier de Lorenzo, R. Queraltó y Francisco García Novo

MATEMÁTICAS. CIENCIA Y EPISTEMOLOGÍA, por Imre Lakatos. Alianza Universal; Madrid, 1981; 380 páginas. Traducción de Diego Ribes. Al fallecer en 1974 Imre Lakatos, un grupo de amigos decidió reunir y editar los diversos ensayos, ya inéditos, ya publicados con anterioridad, del pensador húngaro, nacido en 1922 y exiliado en Gran Bretaña desde 1956. Los *Escritos filosóficos*, al cuidado de John J. Worral y Gregory Currie, aparecieron en dos volúmenes en Cambridge University Press en 1978. Y es el segundo volumen el que ahora ha publicado Alianza Universidad, anunciando para más adelante la versión del primero. Ignoro las razones de esta elección y aunque ambos volúmenes, así como los diversos ensayos en ellos contenidos, se estiman independientes, puede observarse que la intención que los anima vive condicionada por un pensamiento subyacente común: la metodología de los programas de investigación científica, versión dada por Lakatos a la posición de Popper. De manera que alguno de los ensayos aquí reunidos no es otra cosa que una aplicación de tal metodología. Y son los ensayos contenidos en el primero de los volúmenes de la edición inglesa los que discuten y desarrollan el contenido de lo que Lakatos califica como programas de investigación científica. Independiente a esta elección castellana habría que señalar también que los ensayos aquí reunidos no constituyen corpus acabado en el sentido de que Lakatos revisaba de modo continuo sus escritos en función de sus propios cambios de pensamiento, y de aquí que los editores deban suplir con supresiones, añadidos y notas lo escrito por el propio Lakatos. Los editores estructuran los catorce ensayos que componen este volumen en tres apartados: filosofía de la matemática, ensayos críticos y ciencia y educación. Aquí, para una breve crítica, y con tratamiento ciertamente desigual, los agrupo en dos bloques: filosofía de la matemática y escritos polémicos.

1. Filosofía de la matemática. Los comprendidos bajo el epígrafe "filosofía de la matemática" son cinco ensayos, de valor muy desigual, aunque quizá constituyan lo más rico del pensamiento de Lakatos en el sentido de aplicar a la matemática las ideas, ciertamente revisadas, de Popper. Y ello porque la matemática ha constituido de siempre el refugio de lo que Lakatos califica de "programa euclídeo", marginando a cualquier tipo de empirismo, sea el positivista, sea el por él llamado empirismo inductivista. Refugio no sólo por la matemática en sí, sino porque los propios teóricos de la filosofía de la ciencia han marginado sus preocupaciones del hacer matemático y han preferido abocarse a las teorías de la física, fundamentalmente, aceptando que la matemática y la lógica constituyen sistemas de carácter analítico que, en todo caso, pueden estimarse como paradigmas de la auténtica organización que pueda establecerse en cualquier sistema de conocimiento.

Es evidente que sólo un fuerte reconocimiento de la matemática posibilita discutir la pertinencia o no tanto del método falsacionista de pruebas y refutaciones, como de cualquier otro método. En este sentido, debo recordar que en *Investigación y Ciencia* [n.º 24, septiembre de 1978] hice una revisión del pensamiento filosófico de Lakatos con motivo de la publicación en castellano de su obra quizá fundamental *Pruebas y refutaciones*. Aquí, cabe agregar que en este volumen se recogen dos ensayos –de 1962 y de 1967– en los cuales se hace una exposición de los distintos "programas" de fundamentos de la matemática que se han propuesto y que, en el decir de Lakatos, se centran en euclídeo, empirista e inductivista –que es el sostenido por él, siguiendo a Popper–. Del euclídeo da un esquema muy pobre apoyado en el método axiomático, pero en una versión clásica, atribuyéndole la idea de que la matemática fundamenta su verdad en unos primeros principios desde los cuales desciende hasta las proposiciones básicas, mientras que el programa empirista sostiene, según Lakatos, que la inyección de la verdad se realiza en sentido inverso. Naturalmente, el propio Lakatos reconoce que su descripción es muy esquemática y salva su posición mediante la distinción –que ya discutí en la crítica antes mencionada– entre reconstrucción racional e historia real, dejando en el aire la consideración de que la propia historia real puede estimarse como una caricatura de la reconstrucción racional. Y tras la crítica a ambos programas viene a sostener que el objetivo de un programa inductivista de la matemática se centra no en el conocer o en inyectar la verdad, sino en conjeturar y en el establecimiento de "falsadores" que permitan la refutación de teorías ya existentes. El problema es, entonces, la búsqueda de estos falsadores que permitan refutar una teoría y, con ello, posibiliten el cambio de unos a otros problemas, de unos a otros temas más importantes.

Como crítica a esta posición me limito a tres notas: En primer lugar, la propia caracterización de falsador no queda muy clara y es tema al que Lakatos no vuelve en sus ensayos, quizá por la dificultad que encontró en el intento; en segundo lugar, no señala criterios objetivos para denominar y caracterizar lo que califica como problema "más importante"; en tercer lugar, y ligado al primero, el reconocimiento de que teorías muy formalizadas –más bien formalizables– como la de grupos carecen de falsadores salvo los lógicos del tipo "p y no-p", por lo que sólo podrán encontrarse éstos en teorías informales para lo cual Lakatos debe forzar su posición y admitir que teorías como las de conjuntos son casi-empíricas y afirmar que más que falsador hay que hablar de hipótesis rival, porque cualquier falsación en matemáticas se queda en mera sugerencia de dicha falsación. Ello es consecuencia de encontrar Lakatos que axiomas como el de constructibilidad de Gödel o el de ordinales inaccesibles de Tarski, de naturaleza conjuntista muy abstracta, hallan su aplicación en campos matemáticos muy concretos, como en la topología algebraica, por lo que realmente no puede hablarse de falsación, sino de contrastabilidad, Ahora bien, una contrastación positiva no significa nada para un programa de investigación que se ligue a la posición de Popper, donde lo que importa no es, precisamente, tal contrastabilidad positiva sino la negativa, la posibilidad de refutación, de falsación negativa. Y lo que, en este terreno, puede mostrar Lakatos es todo lo contrario de dicha

posición metodológica con pruebas y refutaciones, salvo que llegase a admitir que los axiomas no son portavoces de verdad alguna, sino meros delimitadores de campos de juegos y, por ello, cualquier sistema de postulados lo que hace es posibilitar nuevos campos y, consecuentemente, delimitar unos sistemas teóricos de otros; en el caso citado, teorías conjuntistas que admitan o no determinados axiomas del infinito, con sus consecuencias para otras teorías clásicas como las de topología algebraica ligadas a las álgebras cilíndricas, por ejemplo. Desde el punto de vista, no creo que Lakatos haya logrado una aplicación satisfactoria al total de la matemática de su programa de investigación científica.

Programa que, sin embargo, limitado a una parcela de la matemática y, más en concreto, a la historia de tal parcela, puede mostrar una gran potencia, fundamentalmente en el campo de las sugerencias, en las reconstrucciones racionales de dicha parcela. Y es lo que muestra el tercer ensayo de este volumen que recoge una comunicación de 1966 y que Lakatos se negó reiteradamente a publicar aunque continuara revisando el texto, por la gran cantidad de rectificaciones que una historia real le obligaron a llevar a cabo en su reconstrucción racional. El ensayo es un intento de historia del cálculo de Cauchy enfrentado al cálculo de Weierstrass y los procesos de aritmetización del siglo XIX.

Apasionante ensayo "histórico" en el que se enfrentan dos concepciones del análisis infinitesimal, de una teoría. Y más aún, donde se enfrentan realmente dos filosofías historiográficas: la que quiere ver una continuidad y un progreso lineales y la que ve la historia hecha a base de saltos y rupturas de corte epistemológico. Es claro que Lakatos suscribe esta última, aunque los términos que he utilizado anteriormente no sean suyos sino míos. Deseo hacer constar, aquí, una observación: la terminología de Lakatos en cuanto a "programa de investigación científica" se me muestra confusa porque hace referencia, en ocasiones, a todo el hacer matemático o a todo un sistema científico, mientras que en otras ocasiones, cuando los aplica, hace referencia sólo a teorías muy determinadas y concretas. Y este es el caso del análisis infinitesimal. En el fondo, no hay programa de investigación más o menos falsable, sino una teoría muy concreta y determinada, el análisis infinitesimal, ligado a otras teorías y con unas creencias que lo sustentan y unos alcances dados por unos modelos o realizaciones. Y es la potencia de estos modelos quien va a dar la posibilidad o no de modificaciones en la teoría. Modelos construibles o definibles en términos de, o interiores a, la teoría, en cuyo caso ésta aumenta en cuanto al número de sus proposiciones, o que la refuten, en cuyo caso dicha teoría encuentra alguno de sus límites; pero, al hacerlo, posibilita su propia ampliación para abarcar lo que cae fuera de tales límites; ampliación, por supuesto, que implica un cambio en los conceptos de partida y, consecuentemente, en las creencias que los sustentan. Así, las sucesiones no uniformemente convergentes se estiman como "invenciones del diablo", en términos de Abel, porque no pueden construirse en términos de un cuerpo no arquimediano con variables dinámicas pero, al haberse establecido tales sucesiones, se plantea el problema de cómo delimitarlas o integrarlas, y sólo una aritmetización, es decir, el paso a un cuerpo estrictamente arquimediano, posibilita la definición de dichas sucesiones y su manejo. Y es éste, entre otros motivos, el que conduce a la elaboración de una nueva teoría del análisis, apoyada ahora en un cuerpo arquimediano, y, a la vez, al abandono del lenguaje de los indivisibles que servía de base a los trabajos de Cauchy y Abel. Más que programa de investigación, teoría con un haz de creencias subyacentes.

Los dos ensayos restantes de esta parte se centran en un estudio de la prueba matemática, extraído, en el fondo, de *Conjeturas y refutaciones*, y en una descripción y crítica del método del análisis-síntesis que constituye una réplica a un escrito de Hintikka. Son de desigual valor, con algún error del propio Lakatos en la exposición de su método de pruebas y refutaciones. Las notas de carácter histórico real son de gran valor, a pesar de que Lakatos las emplee en mantener sus ataques a las posiciones de los filósofos analíticos y su tesis de que en matemáticas se trata de contrastar en lugar de probar.

2. Escritos polémicos. Los agrupados bajo el título "ensayos críticos" comprenden seis escritos que podrían reducirse a cinco si los editores, con buen criterio, no hubieran separado el 6 del 11. Estos dos artículos se entresacaron de los borradores en los cuales Lakatos esbozaba una crítica a la obra de Toulmin, *La comprensión humana*. Escrito de carácter polémico, al igual que los restantes de esa sección, dedicados a una crítica del inductivismo de Carnap, de las posiciones de Grunbaum. Kneale, Kuhn e incluso del propio Popper en cuanto a ciertas limitaciones que plantea su metodología al enfrentarse con la historiografía. De hecho tales críticas se resumen en el ensayo que los editores han colocado como el 6 y en el cual Lakatos esquematiza los planteamientos que existen acerca de la evaluación de las teorías científicas. El *escepticismo*, y señala a Feyerabend como máximo exponente, en el que, aun admitiendo que la filosofía de la ciencia sea una actividad perfectamente lícita, insiste en que ningún sistema de creencias es superior a otro, por lo que niega la posibilidad de que pueda resolverse el problema de evaluación de teorías rivales; el *demarcacionismo* pretende producir un criterio universal de evaluación que posibilite identificar el progreso científico y, aun reconociendo que no hay reconstrucción histórica sin sesgo, admite que la filosofía de la ciencia se convierta en el "perro guardián" de las normas científicas; Lakatos se incluye en este punto de vista. Finalmente, el *elitismo*, que sostiene la posibilidad de diferenciar la buena de la mala ciencia aunque no exista un código legal objetivo que sirva como criterio universal o explícito de progreso o degeneraciones; esas normas vendrán establecidas por la comunidad de científicos –es lo sostenido por Toulmin– o por la autoridad académica –es lo sostenido por Kuhn o por Polanyi–, lo cual conduce a mantener líneas ligados al sociologismo y psicologismo en los terrenos científicos así como la posibilidad de intromisión de otros poderes fácticos en los mismos. Es claro que Lakatos no comparte el elitismo, al que considera como autoritario, dogmático e historicista, entre otras apreciaciones. Y desde esta posición es desde la que realiza las críticas que se recogen en los distintos ensayos de esta parte, contraponiendo los criterios, que quiere objetivos, de su metodología de pruebas y refutaciones. Al ser escritos de carácter fundamentalmente polémico –e incluso alguno inacabado– creo que lo mejor que puedo hacer es sugerir su lectura por parte del lector, por la riqueza de matizaciones que los mismos encierran.

Los ensayos agrupados bajo el epígrafe "ciencia y educación" son tres escritos. Comienzan con la carta que Lakatos dirigió al director de la London School of Economics, centro que lo acogió tras su huida de Hungría. La fecha de la carta es de 1968, a raíz de los incidentes provocados por los estudiantes. En ella Lakatos parte de una nítida distinción entre lo que califica "libertad de crítica" y lo que califica "poder de to-

ma de decisiones". Y si concede lo primero a los estudiantes, les niega lo segundo. Lakatos se opone a que los alumnos deban ser admitidos en el Senado y en el Consejo de la Escuela –y parece que fue seguida su sugerencia. Es problema muy de actualidad en ámbitos como los españoles, donde ese tema ha sido y está planteado en los distintos estamentos estudiantiles. El segundo ensayo es el más breve y versa acerca de la enseñanza de la historia de la ciencia, para terminar con otro tema como el de la responsabilidad social de la ciencia, en el que Lakatos aboga por el carácter neutral de la ciencia aunque no de los científicos como individuos.

Tres breves ensayos, los últimos, de carácter radicalmente polémico en áreas socio-políticas con opciones muy claras por parte de Lakatos. El lector debería tener presente –para tratar de situar estos escritos en una mejor perspectiva– que el verdadero nombre de Lakatos es Lipschitz, cambiado por el de Molnár cuando pasó a la clandestinidad bajo la ocupación alemana y adoptado el de Lakatos, cuando la liberación que le condujo a estudiar en Moscú y a ocupar algún puesto en el Ministerio de Educación en Hungría. Encarcelado tres años bajo el estalinismo, consiguió salir, definitivamente, de Hungría en 1956.

La traducción de Diego Ribes es correcta con sólo una salvedad: en el lenguaje matemático castellano se maneja "sucesión" y no "secuencia" –en inglés *sequence*–, aunque es este último el término que aparece reiteradamente en el texto: "cualquier secuencia convergente de funciones continuas" (p. 74), "una secuencia de funciones" (n. 26, p. 77 y p. 81...). En cuanto a la edición contiene algunos errores tipográficos a los que Alianza Universidad nos tiene ya, desgraciadamente, acostumbrados. (J. L.)

FILOSOFÍA DE LA FÍSICA, por Mario Bunge. Barcelona, Ariel, 1978. 301 páginas. Traducido por J. L. García Molina y revisión técnica a cargo del autor. (El título original es: *Philosophy of Physics*. Dordrecht (Holanda), D. Reidel Pub. Co., 1973). Nos hallamos ante una obra de marcado interés dentro del campo filosófico, especialmente el epistemológico, y esto por dos razones principales. En primer lugar por su autor, el conocido filósofo de la ciencia Mario Bunge, que profesa en la actualidad al frente del Instituto de Fundamentos y Filosofía de la Ciencia de la Universidad McGill de Montreal, y una de las más destacadas figuras internacionales en su especialidad. En segundo término, por la misma índole de las tesis mantenidas en el libro, que sin duda originarán –ya lo han hecho– crecientes comentarios y reflexiones.

En obras anteriores, por ejemplo *Scientific Research* de 1967 (trad. esp. de M. Sacristán, Barcelona, 1969), era patente la postura crítica de Bunge frente a determinados aspectos, algunos de ellos fundamentales, de las corrientes comúnmente aceptadas en filosofía de la ciencia, lo cual no hace sino acentuarse en el presente trabajo, y, lo que es muy importante, sistematizarse destacadamente. Su posición filosófica fundamental se ha desarrollado posteriormente, de modo general, en el *Treatise on Basic Philosophy* en siete volúmenes, que a partir de 1974 edita D. Reidel, y cuyas partes tercera y cuarta están dedicadas a la ontología bungeana, con la que la obra que se comenta ahora mantiene no pocas relaciones intrínsecas. Esto significa que las ideas de Bunge expuestas en su *Filosofía de la física* ocupan un lugar específico en el conjunto de su labor filosófica. Ya en el prefacio, y siguiendo con su habitual claridad expositiva, el autor relaciona un elenco de los principales problemas que van a constituir la vertebración interna del libro. La referencia de las teorías físicas a lo real existente, el valor cognoscitivo de las mismas en función de la resolución del problema precedente, la relación teoría-experimento y, finalmente, el valor de los conceptos de la física clásica respecto de la física actual. Con un riguroso orden de conjunto, el capítulo primero establece inicialmente dos fases fundamentales: el rechazo de la denominada filosofía standard de la ciencia, que el autor resume en el operacionalismo, y la ineludible necesidad de una nueva filosofía de la física, la cual ha de servir tanto para obtener avances filosóficos como científicos en cierta medida. El filósofo argentino, radicado en Norteamérica desde hace años, destaca la relevante funcionalidad de la filosofía para las ciencias físicas en la situación actual, circunstancia que a nuestro modo de ver es de todo punto acertada si se tiene presente la incompleta epistemología al uso de los físicos, y la necesidad de claridad conceptual para soslayar, en lo posible, ciertas paradojas de algunas teorías fundamentales de la física. Y esto es lo que el autor considera irrealizable sin acometer una crítica de la filosofía standard, aparte de su rechazo por sus ambigüedades específicas.

En el capítulo segundo se realzan dos ideas capitales del núcleo de la obra: la defensa de la axiomatización de las teorías, como instrumento metodológico básico para el análisis, y la posición ontológica realista de Bunge. La física se ocupa de lo real verdaderamente, y no se trata de una creación idealizada cuya única confirmación se halla en la aplicación tecnológica, sino que ésta es posible porque los resultados de la física se refieren ineluctablemente a las estructuras reales del mundo físico. De aquí se derivarán coherentemente algunas consecuencias básicas que tendrán cumplido desarrollo en capítulos posteriores. Así, en el tercero, se rechaza una interpretación empirista radical de las teorías físicas: éstas no son simples colecciones de datos, poseyendo además mayor significación intrínseca sus componentes teoréticos. El intento de eliminación de los elementos no-observacionales o trans-empíricos de las teorías generales son casi una falacia por la misma definición de teoría científica como sistema hipotético-deductivo (p. 63). La actitud filosófica de Bunge es así necesariamente totalista, recordándose que un carácter esencial de la filosofía es la consideración del todo (p. 64), en lo que el autor parece elevarse, aunque sólo sea en sus aspectos formales, a las concepciones clásicas de la filosofía. (No en vano escuché decir en una ocasión al propio autor que Aristóteles era quizás el más grande de los filósofos.)

Con estas bases, en el capítulo cuarto, que es el de mayor trascendencia en cuanto al contenido filosófico del libro, se opta decididamente por el realismo como ontología básica. Se enumeran cuatro posibles posiciones filosóficas sobre el referente de la física: el realismo (ingenuo y crítico), el subjetivismo (todo enunciado físico se refiere antes o después a un objeto mental), la tesis de Copenhague (enunciado físico como un bloque indiferenciado físico-mental) y la tesis dualista (enunciados como parcialmente físicos y parcialmente de objeto mental). No sería posible desarrollar aquí la crítica completa de Bunge para descartar las tres últimas opciones y adoptar el realismo crítico como posición ontológica propia; tan sólo se indicará que se lleva a cabo a través de un triple camino: mediante el examen de las interpretaciones de los símbolos físicos, sobre la posibilidad de formular una teoría en cada una de las cuatro alternativas y en referencia a la teoría de la medición. El resultado es siempre el mismo: la necesidad fundamentada del realismo crítico. El autor resume, seguidamente, lo que entiende bajo esta

noción en cinco puntos básicos. Primero, hay cosas en sí, o sea, objetos cuya existencia no depende del sujeto. Segundo, tales cosas en sí son cognoscibles parcialmente y por sucesivas aproximaciones. Tercero, este conocimiento se alcanza conjuntamente mediante teoría y experimento, y ninguno de ellos aisladamente puede dictar un veredicto absoluto. Cuarto, el conocimiento científico es hipotético, siendo así corregible y perfeccionable por su propia naturaleza. Y quinto, el conocimiento de esas cosas en sí es indirecto y simbólico (p. 109).

Esta categórica opción por una ontología realista crítica puede suponer una clara alternativa a ciertas bases filosóficas actuales, y sea como fuere es importante señalar que la actitud bungeana puede contribuir a una nueva síntesis epistemológica acerca de la naturaleza última de los resultados y teorías fundamentales de la física, en tanto que aquí se afirma sin paliativos que aquellos se refieren sin matizaciones subjetivistas o instrumentalistas a lo que *es* la realidad física. Sin duda, la tarea que se ha propuesto el autor es de una excepcional envergadura, lo cual no parece ocultársele puesto que en el quinto capítulo reinterpreta la mecánica cuántica en términos realistas, completado en el siguiente con la defensa de una interpretación "literal y objetiva", proporcionándose incluso nuevos conceptos para nombrar las realidades físico-cuánticas. "Deseamos que la ciencia se ocupe de lo que las cosas son", subraya Bunge (p. 147) sin reticencia alguna; ahora bien, ¿es que la ciencia no se refiere a la estructura de lo real, o es que lo hace insuficientemente por razones de muy diversa índole? Sería mejor entender esta afirmación en el sentido de que la comprensión interpretativa de los resultados y teorías científicas se reformulen a fin de nombrar a lo real físico como tal y se eviten posibles ambigüedades. Finalmente, en el capítulo séptimo se defiende la necesidad de una metafísica en la ciencia (p. 170), lo cual no es sino un corolario de todo lo anterior. No hay reparos en adoptar el término metafísica, lo único que se requiere es que se encamine en su dirección adecuada, acorde con unas razones suficientemente fundadas, pues "en todo caso no hay modo de evitar la metafísica" (ibíd.). El objeto básico de tal investigación será descubrir, reunir, sistematizar y realizar un escrutinio crítico de las hipótesis metafísicas presupuestas en la ciencia física (ibíd.). Este sería el programa con su guía heurística fundamental en el realismo crítico, pero todavía no se desarrolla en el libro de forma acabada la metodología adecuada para el mismo, salvo en algunos aspectos por demás importantes. (En su sistematización y aplicación este punto remite al *Treatise on Basic Philosophy*.)

En este aspecto, los capítulos séptimo y octavo se dedican al instrumento básico que el autor establece para los fines enunciados, es decir, la nueva comprensión de las teorías y la filosofía realista de la física. Se trata de la axiomatización. Conviene destacar que Bunge no considera la axiomatización como un fin en sí mismo, sino como camino para la clarificación, ordenación y rigor de las teorías, y procede así a su examen detallado y confrontación mutua. El posible rechazo de la axiomática por los físicos lo contempla como una dificultad práctica importante y se esfuerza en analizar sus ventajas y destruir los diversos "prejuicios" que el científico expondrá ante tal instrumento metodológico. Se formula incluso la concepción de axioma remitiéndose a un sentido etimológico muy amplio –solicitud, petición de análisis–, a fin de eliminar el posible escándalo del físico ante la tesis de la relatividad de los axiomas, o también, desde la posición opuesta, la acusación de autoritarismo al llamar axioma a las hipótesis de partida. De hecho, para Bunge, axiomatizar una teoría es establecer una serie de suposiciones iniciales que caractericen suficientemente los conceptos básicos y generen todos los enunciados o fórmulas standard de la teoría. A su vez, se han de respetar cinco propiedades en la axiomática física: coherencia formal, completud deductiva, completud de los primitivos, independencia de los mismos e independencia de los postulados. Mostrando las ventajas de su propuesta, Bunge ejemplifica con la teoría de redes y la teoría clásica de la gravitación, para después obtener un catálogo de técnicas de axiomatización, sin pretensiones exclusivistas por supuesto. No obstante el valor concedido a la axiomatización, el autor propone una síntesis de ésta con el tratamiento histórico y el heurístico (o sea, el habitual en los trabajos de física) de las teorías. A nuestro juicio hubiera sido deseable un detenimiento mayor en la naturaleza de estos dos últimos, pues se obtiene la impresión de que la axiomática se establece, por la reiteración de su importancia, como algo exclusivo, aún a pesar de las aplicaciones a la enseñanza que se señalan al término del capítulo octavo. Se podría alegar que, puesto que el tratamiento institucionalizado es el heurístico, no sería necesario acentuar su importancia, mientras que es procedente incidir en la axiomatización. No obstante, queda la impresión, por la extensión concedida al método axiomático y por la relativa al tratamiento heurístico, que esta manera habitual de enfoque se considera excesiva en una analítica de las teorías.

El capítulo noveno se dedica al examen de las relaciones entre teorías, en donde se destaca primeramente lo poco realizado hasta el momento en este campo si se compara con otros más clásicos o más atendidos. Se analiza el tema de la reducción de teorías, en el cual Bunge señala la problematicidad de intentos tales como la reducibilidad de la mecánica clásica a la cuántica, afirmando que no es posible establecer conclusiones definitorias como muchos pretenden. Igualmente, tampoco sería procedente considerar ciertas teorías como "límite" superior o inferior de otras; por el contrario, se conciben las teorías fundamentales de la física con un grado de independencia intrínseca entre ellas, y en todo caso se apunta hacia un análisis más estricto –con el instrumento axiomático– de cada teoría a fin de poder afirmar algunos niveles o puntos de reducibilidad. La crítica bungeana, ya conocida, a un inductivismo o un empirismo extremados en filosofía de la ciencia se hace patente a lo largo de este capítulo, pues se reiteran las tesis de que las contrastaciones empíricas no vienen dictadas únicamente por la propia teoría en examen, sino que ésta necesita la ayuda de otras teorías. Asimismo se destaca el valor de las contrastaciones teóricas o de compatibilidad general con el corpus de conocimiento anterior, lo cual complementa el rechazo procedente del empirismo, y finalmente se reafirma la concepción de teoría como sistema hipotético-deductivo frente al "absurdo credo" de la síntesis inductiva (p. 250).

De acuerdo con esta posición esencial, mantenida con claridad y rigor en toda la obra, se articula el último capítulo dedicado a la relación entre experimento y teoría. Bunge destaca el carácter teórico que atraviesa cualquier interpretación de los datos empíricos, rechazando un poder autointerpretativo de los mismos, lo que no es sino consecuencia lógica de sus tesis básicas. El autor resume su intención al señalar "la íntima trabazón entre teoría y experiencia, la cual refuta la presunción de que cualquiera de los dos polos es dominante" (p. 251). Al llevar a cabo esta suerte ideal de equilibrio se dedican las cuatro clases de "baterías de pruebas"

de una teoría: metateóricas, interteóricas, filosóficas y empíricas, señalándose la prioridad de las tres primeras, o sea, las contrastaciones no empíricas. No pretente el autor con esto desmerecer el nivel empírico, sino hacer ver la elevada dosis teorética de la física actual. Punto conflictivo puede ser sin duda la denominada contrastación filosófica, pues se trata a fin de cuentas de cuestiones de principios filosóficos, donde la opción pudiera quedar menos clarificada por la abstracción inherente a dicho análisis. Ahora bien, la necesidad de este examen crítico es un hecho, y Bunge recuerda que toda investigación científica está penetrada por ideas filosóficas, independientemente de la autoconsciencia de las mismas. En definitiva, se persigue una unión significativa y fructífera entre ciencia y filosofía que reduzca la separación consumada a lo largo de varios siglos, y que en el momento actual parece evidenciarse como un error cuyas consecuencias epistemológicos se han tornado de una importante gravedad.

Es ésta, por último, la idea directriz que anima la presente obra de uno de los más significativos filósofos de la ciencia en la actualidad. Lo más destacable de ella es que Bunge propone todo un programa concreto y *fundamentado* a fin de sistematizar coherentemente el camino hacia sus objetivos. En otros términos, no se trata de una expresión de propósitos, sino que se acomete la investigación desde sus principios sin eludir cuestiones filosóficas espinosas. Lógicamente, no se excluyen por esto algunas reservas críticas a determinadas aseveraciones del autor, que no pueden analizarse aquí, pero se ha de destacar el rigor y la validez de la tarea emprendida. En ese sentido esta obra, así como el *Treatise on Basic Philosophy* y la aparecida últimamente con el título *Epistemología* (Barcelona, Editorial Ariel, 1980), suponen un legítimo y valioso esfuerzo por eliminar prejuicios e ideas cuyo ciclo histórico parece terminar, pudiéndose esperar que la labor propuesta, y en parte ya realizada, adquiera una importante relevancia futura y contribuya, especialmente, a ese acercamiento y trabajo común entre filósofos y científicos que se presenta hoy día como tarea necesaria. (R. Q.)

LA BIOSFERA. ENTRE LA TERMODINÁMICA Y EL JUEGO, por Ramón Margalef. Editorial Omega; Barcelona, 1980; 236 p., 21 figuras. El título de la obra resulta algo desconcertante, pero el propio autor lo presenta como homenaje a un título clásico en la ciencia biológica: *El azar y la necesidad*, de J. Monod. El "azar" es ahora el juego en una perspectiva más amplia que engloba, junto a los fenómenos aleatorios, las estrategias de los jugadores y las reglas. La "necesidad" es la limitación que imponen los intercambios energéticos a nuestro universo y cuyo tratamiento suele hacerse en la termodinámica. El contenido de la obra, fiel al título, aborda la organización de los sistemas ecológicos de la biosfera dentro de su marco energético.

Hay que destacar que el libro no es un manual o una abreviación de la *Ecología* del mismo autor, aparecida en 1974. Se sitúa más bien en la línea de *Perspectives in ecological theory* (1968, publicada su traducción española en 1978) y particularmente *La teoría de la información en Ecología* (Mem. Acad. Cien. Art. Barcelona, 40: 3-22, 1969) y "On certain unifying principles in ecology" (*Amer. Nat.* 97: 357-74, 1963).

Margalef teórico reelabora una "ecología de sistemas" bajo perspectiva física. Pero no trata de sistematizar mucho sus ideas en busca del rigor expositivo, dejándose apresar por el lenguaje sistemico o la formulación matemático como es costumbre en ecología teórica. Por el contrario, todo el libro huye de la formalización, dando preferencia a las expresiones y modelos verbales sobre los otros y presentando el autor sus puntos de vista como opiniones o sugerencias plausibles más que como evidencias científicas absolutas. La redacción en primera persona y las referencias frecuentes a experiencias suyas ayudan a conformar el libro como una pieza provocativa y abierta del pensamiento ecológico.

La Biosfera es un libro de formato manual, pero su extensión no es pequeña; 215 páginas de letra menuda, casi sin ilustraciones, que lo hacen particularmente denso y apretado. El contenido se distribuye en seis capítulos que se articulan en tres partes. La primera (particularmente innovadora) incluye el capítulo 1: *La ecología, las ciencias físicas y el hombre*, que introduce la temática del libro y fija la posición del autor frente a las actuales ecologías de sistemas, modelística y ecología aplicada. El capítulo 2 se titula *El juego en el límite de los sistemas termodinámicos.*

Esta primera parte de *La Biosfera* resulta la más atractiva del libro. Expresando las ideas alrededor del tema central de "información" las enlaza con otras medidas de orden y energía y las asocia a la organización y memoria, interpretando, a la luz de este espectro de ideas, muchos de los sistemas de los que formamos parte, desde las galaxias hasta el lenguaje. La lectura resulta particularmente sugerente y evocadora, abriendo un abanico de perspectivas a la ecología, biología y otras ciencias.

El resto del libro reitera o desarrolla, de alguna manera, los temas de los primeros capítulos; a veces se limita a modificar algo las tesis mantenidas por el autor con anterioridad. *La diversidad* (cap. 3), un tema favorito de Margalef, se presenta con una gran recopilación (poco concluyente) de datos en sistemas muy diferentes. Apenas se detiene en las intrigantes relaciones de diversidad y estabilidad. Esta última, empero, recibe un tratamiento más profundo en el capítulo 4: *Interacciones entre los constituyentes del ecosistema.* Utiliza como hilo conductor la información y precisa el grado de asimetría en las relaciones y las posibilidades de subdivisión en estructuras (subsistemas). Semejante planteamiento le permite una buena perspectiva del tratamiento de las interacciones entre poblaciones y una crítica de los modelos básicos depredador-presa y competencia; y justifica la necesidad de considerar sistemas no binarios (ternarios o de más elementos) y establecer en la formulación un enfoque que pondere la ganancia diferente de información que se da entre uno y otros elementos interactuantes.

En la tercera parte, el capítulo 5, que lleva por título *Estabilidad y cambio en sistemas complejos*, una compendiada ecología de sistemas, aborda la mayor parte de los principios sistémicos. El sexto, sobre *La organización de la Biosfera*, le da pie para aplicar algunos de los principios discutidos en los anteriores a la organización de ecosistemas, evolución, filogenia y sucesión. Completa el libro su extensa bibliografía (305 entradas), compuesta de citas y obras interesantes no aludidas en el texto. Predominan los trabajos de ecología teórica y son muy escasas las referencias a la teoría general de sistemas. Un índice alfabético facilita la utilización de la obra.

La explosión de publicaciones ecológicas a la que estamos asistiendo, la acumulación de datos, el desarrollo de modelos ¿nos están haciendo avanzar en ecología? La sensación personal es que poco. La tesis de Margalef en *La Biosfera* es que los avances podrán darse, pero en otras direcciones distintas, particularmente en el estudio físico de los sistemas biológicos y en el discernimiento de las relaciones hombre-biosfera.

A decir verdad, el desarrollo de la física ha resultado insuficiente hasta hace poco, proporcionando instrumentos muy rudimentarios para los estudios

ecológicos, que sólo permitían la medida de rasgos macroscópicos de los sistemas ecológicos, pero apenas nada de los que constituyen la regulación, organización, creación y transferencia de información. El desarrollo de una termodinámica de sistemas abiertos en equilibrio dinámico constituye un gran paso en esta dirección. Margalef expresa esta necesidad de proyección física de la teoría ecológica, en la primera página del prólogo, "considerando a la biosfera en el marco de una teoría física muy amplia". Su preocupación emerge una y otra vez cuando habla de energía y, especialmente, en sus esfuerzos por asociar el intercambio de información al de la energía y las relaciones de uno y otro con la organización de un sistema.

Hay un aspecto importante donde la física ha conseguido profundizar y que podría seguirse en *La Biosfera*: un tratamiento más discriminado del significado físico de los procesos temporales. El capítulo 2 (reversibilidad de las interacciones, memoria, cambio) trata del tiempo, o está presente como condicionante temporal de los procesos. Pero el significado heterogéneo que le da la física actual, tal como lo presenta por ejemplo Prigogine (*From being to becoming*, Freeman 1980), resulta sugerente para el ecólogo.

Aunque *La Biosfera* es claramente un libro de teoría ecológica, posee un interesante moderador que le evita perderse en la especulación gratuita. Margalef parte de la realidad inmediata y se ve "forzado" a explorar los principios teóricos en busca de soluciones a los problemas concretos que la investigación ecológica plantea. Las referencias a datos, fenómenos y ejemplos inmediatos son constantes en el texto, acompañadas de un cierto desdén por los desarrollos teóricos "per se" y los modelos en particular. Por esto no podía faltar el capítulo sexto, dedicado a la biosfera, donde el autor disfruta aplicando a nivel concreto las consecuencias de los principios teóricos elaborados en los capítulos anteriores.

El hombre como elemento activo en los ecosistemas está presente en todo el libro, no de forma tangencial, sino con el convencimiento (que expresa el autor ya en el prólogo) de que el estudio de tales interacciones constituye una línea maestra para el desarrollo de la ecología. El hombre es objeto de estudio ecológico, mas no en el sentido de la ecología humana, sino, sobre todo, a través de la sociedad industrial siguiendo la idea de H. T. Odum en su *Environment, power and society*. En *La Biosfera* este análisis queda sólo esbozado; es claro que el autor no deseaba insistir en aspectos sociológicos que hubieran pesado excesivamente en el libro, alejándolo de su objetivo focal de la biosfera.

Con todo se percibe una cierta desgana sobre la aplicación ecológica, en especial en sus vertientes de impacto ambiental y modelística. ¿Por qué? Precisamente uno de los logros más interesantes de la ecología como disciplina ha sido permeabilizar la sociedad industrial despertando su sensibilidad a las intervenciones en el entorno. El ecologismo y las evaluaciones del impacto ambiental son otras tantas expresiones de esta preocupación social y permiten la aplicación de conocimientos ecológicos a la gestión de recursos. Cierto que la ecología no es ecologismo, ni los ecólogos pueden confundirse con los ecologistas. Pero, como en el impacto ambiental, la ecología es fuente de inspiración y apoyatura científica de ambos. La modelística, por su parte, es sólo una herramienta para el ecólogo y bien limitada en sus alcances. Sin embargo, dentro de estas acotaciones, tan válida como el análisis o la estadística y especialmente apropiada como instrumento para abordar relaciones complejas.

La crítica común a impactos, modelística y ecologismo desde el punto de vista ecológico tiene que ver, probablemente, con la facilidad con que caen en tratamientos superficiales de los problemas, generando resultados triviales y de nulo valor científico. Pero el problema no sería entonces de índole estrictamente científica, sino sociológico, y, en cierta medida, ajeno a la ciencia ecológica.

La obra resulta de lectura laboriosa, por manejar conceptos abstractos en forma densa y poco formalizada, incrustada de sugerencias y atisbos difíciles de seguir. Con todo, resulta muy sugestiva y rica, alumbrando parcelas oscuras de la biosfera, proponiendo rutas a seguir y rutas a abandonar en su exploración. El interés no queda limitado a los ecólogos: otros científicos que abordan sistemas complejos como geólogos, químicos, sociólogos y por supuesto biólogos pueden encontrar informaciones, isomorfismos con su propia disciplina y muchos puntos de atracción e inspiración.

Desde el punto de vista editorial, se trata de un libro cuidado, donde es difícil encontrar erratas en el texto o la bibliografía. Resulta estimulante su publicación como primicia en español, sin tener que esperar muchos años la traducción de su versión inglesa. Todos mis elogios para esta publicación de Omega, con la esperanza de que el ejemplo se multiplique. (F. G. N.)

Bibliografía

Los lectores interesados en una mayor profundización de los temas expuestos pueden consultar los trabajos siguientes:

INTENSIFICADOR DE IMAGENES POR MICROCANALES

CHANNEL ELECTRON MULTIPLIERS. Tres números especiales de *Acta Electronica*: vol. 14, n.º 1, enero, 1971; vol. 14, n.º 2, abril, 1971; vol. 16, n.º 1, enero, 1973.

SECONDARY ELECTRON EMISSION MULTIPLIERS AS PARTICLE DETECTORS. W. E. Baumgartner y W. K. Huber en *Journal of Physics E: Scientific Instruments*, vol. 9, n.º 5, págs. 321-330; mayo, 1976.

MICROCHANNEL PLATE DETECTORS. Joseph L. Wiza en *Nuclear Instruments and Methods*, vol. 162, págs. 587-601; 1979.

HIPOCRATES

LA CURACIÓN POR LA PALABRA EN LA ANTIGÜEDAD CLÁSICA. P. Laín. Revista de Occidente; Madrid, 1958.

ENFERMEDAD Y PECADO. P. Laín. Editorial Toray. Barcelona, 1961.

THERAPEIA. Luis Gil. Ediciones Guadarrama; Madrid, 1969.

HIPÓCRATES Y LA NOSOLOGÍA HIPOCRÁTICA. E. Vintró. Editorial Ariel; Barcelona, 1973.

TRACTATS MÈDICS, Hipòcrates. Traducción catalana de Alsina-Vintró. "Fundació Bernat Metge", n.os 184 y 196; Barcelona, 1972-1976.

LA MEDICINA HIPOCRÁTICA. Dirigido por P. Laín. Editado por el C.S.I.C.; Madrid, 1976.

REPARACION INDUCIBLE DEL ADN

recA PROTEIN PROMOTES HOMOLOGOUS-PAIRING AND STRAND-EXCHANGE REACTIONS BETWEEN DUPLEX DNA MOLECULES. Stephen C. West, Era Cassuto y Paul Howard-Flanders en *Proceedings of the National Academy of Sciences of the United States of America*, vol. 78, n.º 4, págs. 2100-2104; abril, 1981.

PURIFIED *LEXA* PROTEIN IS A REPRESSOR OF THE *RECA* AND *LEXA* GENES. John W. Little, David W. Mount y Celeste R. Yanisch-Perron en *Proceedings of the National Academy of Sciences of the United States of America*, vol. 78, n.º 7, págs. 4199-4203; julio, 1981.

ANILLOS DEL SISTEMA SOLAR

JUPITER'S RINGS. Tobias Owen, G. Edward Danielson, Allan F. Cook, Candice Hansen, Virginia L. Hall y Thomas C. Duxbury en *Nature*, vol. 281, n.º 5731, págs. 442-446; 11 de octubre de 1979.

ENCOUNTER WITH SATURN: VOYAGER 1 IMAGING SCIENCE RESULTS. Bradford A. Smith, Lawrence Soderblom, Reta Beebe, Joseph Boyce, Geoffrey Briggs, Anne Bunker, Stewart A. Collins, Candice J. Hansen, Torrence V. Johnson, Jim L. Mitchell, Richard J. Terrile, Michael Carr, Allen F. Cook II, Jeffrey Cuzzi, James B. Pollack, G. Edward Danielson, Andrew Ingersoll, Merton E. Davies, Garry E. Hunt, Arnold Masursky, Eugene Shoemaker, David Morrison, Tobias Owen, Carl Sagan, Joseph Veverka, Robert Strom y Verner E. Suomi en *Science*, vol. 212, n.º 4491, págs. 163-191; 10 de abril de 1981.

LAS FASES FLUIDAS DE LA MATERIA

STATES OF MATTER. David L. Goodstein. Prentice-Hall, Inc., 1975.

THEORY OF SIMPLE LIQUIDS. Jean Pierre Hansen e Ian R. McDonald. Academic Press, 1976.

WHAT IS "LIQUID"? UNDERSTANDING THE STATES OF MATTER. J. A. Barker y D. Henderson en *Reviews of Modern Physics*, vol. 48, n.º 4, págs. 587-671; octubre 1976.

PATRONES DE COLORACION DE LAS MARIPOSAS DIURNAS Y NOCTURNAS

WING PATTERN FORMATION IN LEPIDOPTERA: A MODEL. H. Frederik Nijhout en *The Journal of Experimental Zoology*, vol. 206, n.º 2, págs. 119-136; noviembre 1978.

ONTOGENY OF THE COLOR PATTERN ON THE WINGS OF *PRECIS COENIA* (LEPIDOPTERA: NYMPHALIDAE). H. Frederik Nijhout en *Developmental Biology*, vol. 80, n.º 2, págs. 275-288; diciembre, 1980.

PATTERN FORMATION ON LEPIDOPTERAN WINGS: DETERMINATION OF AN EYESPOT. H. Frederik Nijhout en *Developmental Biology*, vol. 80, n.º 2, págs. 267-274; diciembre, 1980.

ESTABLECIMIENTO BALLENERO VASCO DEL SIGLO XVI EN EL LABRADOR

BALLENEROS CÁNTABROS. Rafael González Echegaray. Editado por la Institución Cultural de Cantabria (C.S.I.C.) y la Caja de Ahorros de Santander. Santander, 1978.

LOS BALLENEROS VASCOS EN CANADA ENTRE CARTIER Y CHAMPLAIN. Selma Huxley Brakham. Real Sociedad Vascongada de los Amigos del País, 1979.

DIVERS FIND SUNKEN BASQUE GALLEON IN LABRADOR. Selma Huxley Barkham and Robert Grenier en *Canadian Geographical Journal*, vol. 97, n.º 3, págs. 60-63; julio, 1979.

EL REVESTIMIENTO DEL INTESTINO DELGADO

TRANSPORT ACROSS THE INTESTINAL MUCOSAL CELL: HIERARCHIES OF FUNCTION. D. S. Parsons y C. A. R. Boyd en *International Review of Cytology*. vol. 32, págs. 209-255; 1972.

THE INTESTINAL CELL SURFACE: PROPERTIES OF NORMAL, UNDIFFERENTIATED, AND MALIGNANT CELLS. Kurt J. Isselbacher en *The Harvey Lectures*, vol. 69, págs. 197-221; 1973-74.

FIFTEEN YEARS OF STRUGGLE WITH THE BRUSH BORDER. Robert K. Crane en *Intestinal Absorption and Malabsorption*, dirigido por T. Z. Csáky. Raven Press, 1975.

TEMAS METAMAGICOS

SIMPLE MATHEMATICAL MODELS WITH VERY COMPLICATED DYNAMICS. Robert M. May en *Nature*, vol. 261, n.º 5560, págs. 459-467; 10 de junio de 1976.

ITERATED MAPS ON THE INTERVAL AS DYNAMICAL SYSTEMS. Pierre Collet y Jean-Pierre Eckmann. Birkhauser Boston, Inc., 1980.

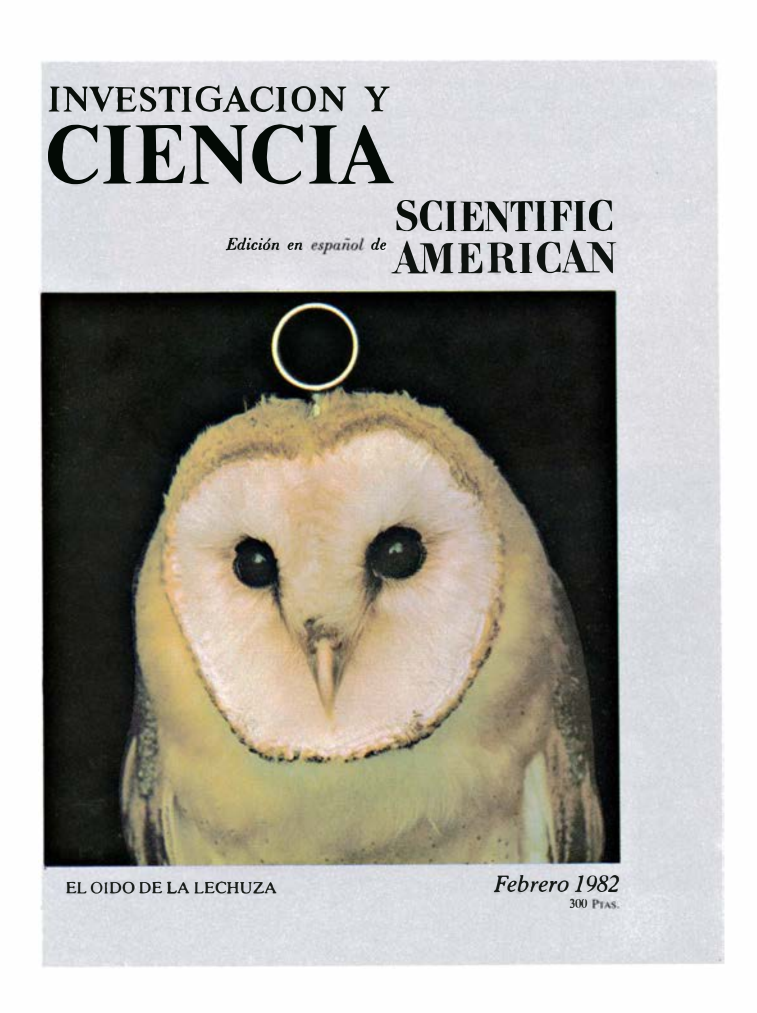
UNIVERSAL BEHAVIOR IN NONLINEAR SYSTEMS. Mitchell J. Feigenbaum en *Los Alamos Science*, págs. 4-27; verano de 1980.

TALLER Y LABORATORIO

PINHOLE IMAGERY. M. Young en *American Journal of Physics*, vol. 40, n.º 5, págs. 715-720; mayo, 1972.

FIELD-WIDENED PINHOLE CAMERA. John M. Franke en *Applied Optics*, vol. 18, n.º 17, págs. 2913-2914; septiembre, 1979.

Seguiremos explorando los campos del conocimiento



EL OIDO DE LA LECHUZA, por Eric I. Knudsen
El ave aprovecha las diferencias entre los sonidos que captan sus oídos izquierdo y derecho para localizar ratones en la oscuridad.

FIBRINOGENO Y FIBRINA, por Russell F. Doolittle
El fibrinógeno, una proteína plasmática, se transforma en fibrina y produce un coágulo sanguíneo. Al cabo de un tiempo, el coágulo se desintegra. Se explican ahora ambos procesos a partir de la estructura de las proteínas implicadas.

EL PRINCIPIO ANTROPICO, por George Gale
Ciertas condiciones, como la temperatura, favorecieron la aparición de vida en la tierra. El principio antrópico sostiene lo contrario: la presencia de vida podría "explicar" las condiciones.

ARMAS LASER, por Kosta Tsipis
¿Podría defenderse un territorio del ataque de misiles mediante láseres en órbita? Los obstáculos tecnológicos son insuperables; además, las contramedidas más simples los anularían.

NAVEGACION MAGNETICA EN LAS BACTERIAS, por Richard P. Blakemore y Richard B. Frankel
Ciertas bacterias acuáticas son magnetotácticas: poseen pequeñas brújulas internas que las orientan en el campo magnético terrestre.

JUPITER Y SATURNO, por Andrew P. Ingersoll
Dos modelos rivales tratan de describir los compañeros gigantes del Sol. En un modelo, los vientos están confinados a una delgada película de la superficie; en el otro, los vientos se extienden a través de las profundidades fluidas de cada planeta.

LIMITACIONES DE LA FIJACION BIOLOGICA DE NITROGENO, por Eulogio J. Bedmar y José Olivares
Hay factores limitantes en la fijación biológica de nitrógeno que representan pérdidas importantes de energía, cuya superación puede contribuir a incrementar de forma considerable la cantidad de proteína disponible.

ALGEBRA POR ORDENADOR, por Richard Pavelle, Michael Rothstein y John Fitch
Un ordenador puede tratar símbolos y números. Nuevos algoritmos de aplicación general pueden realizar múltiples rutinas matemáticas y resolver problemas hasta ahora inabordables.

